E-mail: meliot@phare.normalesup.org

Abstract In this paper, we review the representation theory of the infinite symmetric group, and we extend the works of Kerov and Vershik by proving that the irreducible characters of $\mathfrak{S}_\infty$ always satisfy a central limit theorem. Hence, for any point of the Thoma simplex, the corresponding measures on the levels $\mathfrak{Y}_n$ of the Young graph have a property of gaussian concentration. By using the Robinson-Schensted-Knuth algorithm and the theory of Pitman operators, we relate these results to the properties of certain random permutations obtained by riffle shuffles, and to the behaviour of random walks conditioned to stay in a Weyl chamber.

A central limit theorem for the characters of the infinite symmetric group and of the infinite Hecke algebra

Pierre-Loïc Méliot

For $n \geq 0$, we denote by $\mathfrak{S}_n$ the **symmetric group** of order n, that is to say the group of permutations of $[\![1, n]\!] = \{1, 2, \ldots, n\}$. We also denote by $\mathfrak{S}_\infty$ the inductive limit $\bigcup_{n=0}^\infty \uparrow \mathfrak{S}_n$. The representation theory of $\mathfrak{S}_\infty$ is quite difficult, because it is a wild group, *i.e.*, an infinite group with factor representations of type II or III. However, the classification of the finite factor representations of $\mathfrak{S}_\infty$ is entirely known since the works of E. Thoma (*cf.* [Tho64]). Hence, these representations are labelled by the pairs $\omega = (\alpha, \beta)$ of non-increasing sequences of non-negative real numbers

$$\alpha = (\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n \geq \cdots) \qquad ; \qquad \beta = (\beta_1 \geq \beta_2 \geq \cdots \geq \beta_n \geq \cdots)$$

such that $\sum_{i=1}^\infty \alpha_i + \sum_{i=1}^\infty \beta_i \leq 1$. Moreover, if χ^ω is the normalized character of $\mathfrak{S}_\infty$ corresponding to ω, and if $\sigma \in \mathfrak{S}_\infty$ is a product of disjoint cycles of lengths $l_1, \ldots, l_r \geq 2$, then

$$\chi^\omega(\sigma) = \prod_{j=1}^r \left(\sum_{i=1}^\infty (\alpha_i)^{l_j} + (-1)^{l_j - 1} \sum_{i=1}^\infty (\beta_i)^{l_j} \right).$$

Now, if one considers a restriction $\chi^\omega_{|\mathfrak{S}_n}$ of χ^ω to a finite symmetric group, it is not in general an irreducible character of $\mathfrak{S}_n$, and one can expand it as a linear combination

$$\chi^\omega_{|\mathfrak{S}_n} = \sum_{\lambda \in \mathfrak{Y}_n} \mathbb{P}_{n,\omega}[\lambda] \, \chi^\lambda$$

of the normalized irreducible characters of $\mathfrak{S}_n$. The sum runs over the set $\mathfrak{Y}_n$ of (isomorphism classes of) irreducible representations of $\mathfrak{S}_n$, which happen

The Gaspard–Monge Institute of electronics and computer science, University of Marne-La-Vallée Paris-Est, 77454 Marne-la-Vallée Cedex 2, France

to correspond bijectively to the **integer partitions** of size n, that is to say the non-decreasing sequences of positive integers that sum to n:

$$\mathfrak{Y}_n = \left\{ \lambda = (\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r) \;\middle|\; |\lambda| = \sum_{i=1}^{r} \lambda_i = n \right\}.$$

Hence, for any pair $\omega = (\alpha, \beta)$ and any integer n, one obtains a probability measure $\mathbb{P}_{n,\omega}$ on $\mathfrak{Y}_n$. It has been shown by Kerov and Vershik (see [KV81a, KV81b]) that these systems of probability measures, called coherent systems, satisfy a law of large numbers. Thus, the irreducible characters χ^λ (viewed as random functions on the symmetric groups) and the renormalized coordinates of the random partitions λ under $\mathbb{P}_{n,\omega}$ both converge in probability. In this paper, we will refine this result by showing that the convergence is always gaussian, see our main Theorems 17, 18 and 25.

The article is organized as follows. In the first part of the paper, we recall the results of E. Thoma, using extensively the formalism of symmetric functions. In Section 2, we explain how coherent systems of measures are related to models of random permutations obtained by generalized riffle shuffles; here we use well-known arguments from the theory of symmetric and quasi-symmetric functions. In §3, we present a "method of non-commutative moments" that is due to Kerov and Olshanski (see [KO94, IO02, Śni06b]), and that is extremely useful in the setting of asymptotic representation theory of the symmetric groups.

We apply this method in Sections 4 and 5 to reprove the law of large numbers of Kerov and Vershik and to establish our central limit theorems. The arguments are exactly those that the author used with V. Féray in [FM10] for the asymptotic study of the q-Plancherel measures; in our general setting, they shall seem somewhat simpler. However, when looking at the fluctuations of the rows and columns of the random partitions taken with respect to the probability $\mathbb{P}_{n,\omega}$, new difficulties arise when some coordinates of the parameter ω in the Thoma simplex are equal. We address these difficulties in Section 6 by using and generalizing arguments due to N. O'Connell (*cf.* [O'C03], see also [Lit95, OY02, BBO05, Bia09]); hence, we relate a random walk on $\mathfrak{Y} = \bigsqcup_{n \in \mathbb{N}} \mathfrak{Y}_n$ whose marginale laws are the $\mathbb{P}_{n,\omega}$'s, to a multidimensional random walk conditioned to stay in a Weyl chamber. This link ensures that the fluctuations of rows or columns corresponding to the same parameter are the same as those of the eigenvalues of a gaussian hermitian matrix. Incidentally, the link between random partitions and random walks provides a new proof of some results of Section 5, in particular Theorem 19.

Finally, in §7, we explain how our results behave with respect to the quantization of the symmetric group, that is to say that we replace $\mathfrak{S}_\infty$ by its Hecke algebra $\mathscr{H}(\mathfrak{S}_\infty)$. In particular, we perform all the computations in the

case when the irreducible character of $\mathscr{H}_q(\mathfrak{S}_\infty)$ is the so-called Jones-Ocneanu trace ([Jon87]) of parameter $z = -(1-q)(1-t)$, thereby providing an asymptotic study of a (q,t)-deformation of the Plancherel measures of the symmetric groups.

We have tried to make this paper accessible to a wide audience, and self-contained. In particular, we will recall many well-known results on symmetric groups, symmetric functions and (free) quasi-symmetric functions.

1 Representation theory of the infinite symmetric group

In this first paragraph, we recall the representation theory of the infinite symmetric group, and a basic tool that we will use throughout the paper, namely, the Hopf algebra of symmetric functions (later, we shall also need the algebras of quasi-symmetric and free quasi-symmetric functions).

1.1 Representation theory of the symmetric groups $\mathfrak{S}_n$

To begin with, let us recall the representation theory of the finite symmetric groups $\mathfrak{S}_n = \mathfrak{S}(\llbracket 1, n \rrbracket)$. If $\sigma \in \mathfrak{S}_n$ is a permutation, it can be written uniquely as a product of disjoint cycles; for example,

$$41728635 = (1, 4, 2)(3, 7)(5, 8)(6).$$

Then, two permutations are conjugated in $\mathfrak{S}_n$ if and only if the lengths or their cycles form the same integer partition (here, $(3, 2, 2, 1)$). Thus, the conjugacy classes of $\mathfrak{S}_n$ are labelled by the elements of $\mathfrak{Y}_n$. The irreducible representations of $\mathfrak{S}_n$ are also labelled by integer partitions of size n. Indeed, if $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_r)$, let us denote

$$\Delta^\lambda(x_1, \ldots, x_r) = \prod_{i=1}^r \Delta(x_{\lambda_1 + \cdots + \lambda_{i-1} + 1}, \ldots, x_{\lambda_1 + \cdots + \lambda_i})$$

$$= \prod_{i=1}^r \left(\prod_{\lambda_1 + \cdots + \lambda_{i-1} + 1 \leq j < k \leq \lambda_1 + \cdots + \lambda_i} x_j - x_k \right).$$

The symmetric group acts (on the right) on $\mathbb{C}[x_1, \ldots, x_n]$ by permutation of the variables, and the complex vector space V generated by the $(\Delta^\lambda) \cdot \sigma$

with $\sigma \in \mathfrak{S}_n$ happens to be an irreducible module over $\mathfrak{S}_n$; moreover, all the isomorphism classes of irreducibles can be obtained this way (*cf.* [JK81]).

If λ is a partition of size n, we represent it by its **Young diagram**: this is the array of n boxes with λ_1 boxes on the first row, λ_2 boxes on the second row, *etc.*

$$(5,4,2,2) \quad \leftrightarrow \quad \text{(Young diagram)}.$$

The conjugate of a partition λ is the partition λ' obtained by symmetrizing its Young diagram with respect to the first diagonal; for instance, $(5,4,2,2)' = (4,4,2,2,1)$. In the following, we shall denote by V^λ the irreducible module over $\mathfrak{S}_n$ associated to the polynomial $\Delta^{\lambda'}$; this is the usual convention for encoding the values of the irreducible characters in the algebra of symmetric functions, see §1.2. The two one-dimensional representations of $\mathfrak{S}_n$ correspond to the partitions (n) and $(1,1,\ldots,1) = 1^n$: the first one gives the trivial representation and the second one is the sign representation.

The Young diagrams yield a simple description of the branching rules of the irreducible modules of $\mathfrak{S}_n$ and $\mathfrak{S}_{n+1}$. If $\lambda \in \mathfrak{Y}_n$ and $\Lambda \in \mathfrak{Y}_{n+1}$, we denote $\lambda \nearrow \Lambda$ if Λ can be obtained from λ by adding a single box to the Young diagram.

Example. If $\lambda = (5,4,2,2)$, the partitions Λ of size $14 = 13 + 1$ such that $\lambda \nearrow \Lambda$ are $(6,4,2,2)$, $(5,5,2,2)$, $(5,4,3,2)$ and $(5,4,2,2,1)$.

The **branching rules** of the symmetric groups are the following (see [OV04], [Zel81]):

$$\forall \lambda \in \mathfrak{Y}_n, \ \mathrm{Ind}_{\mathfrak{S}_n}^{\mathfrak{S}_{n+1}}(V^\lambda) = \sum_{\Lambda \in \mathfrak{Y}_{n+1} \ | \ \lambda \nearrow \Lambda} V^\Lambda$$

$$\forall \Lambda \in \mathfrak{Y}_{n+1}, \ \mathrm{Res}_{\mathfrak{S}_n}^{\mathfrak{S}_{n+1}}(V^\Lambda) = \sum_{\lambda \in \mathfrak{Y}_n \ | \ \lambda \nearrow \Lambda} V^\lambda.$$

Since the restriction of representations preserves the dimension, by applying n times the second branching rule, one sees that

$$\dim \lambda = \dim V^\lambda = \mathrm{card}\,\{\text{standard tableaux of shape } \lambda\},$$

where a standard tableau of shape λ is a numbering of the boxes of the Young diagram of λ that is strictly increasing along the rows and along the columns.

Example. The five standard tableaux of shape $(3,2)$ are

$$\begin{array}{cc} 4 & 5 \\ 1 & 2 & 3 \end{array} \ ; \quad \begin{array}{cc} 3 & 5 \\ 1 & 2 & 4 \end{array} \ ; \quad \begin{array}{cc} 2 & 5 \\ 1 & 3 & 4 \end{array} \ ; \quad \begin{array}{cc} 3 & 4 \\ 1 & 2 & 5 \end{array} \ ; \quad \begin{array}{cc} 2 & 4 \\ 1 & 3 & 5 \end{array}.$$

Hence, $\dim(3,2) = 5$.

There is also a "closed" formula for $\dim \lambda$ known as the **hook length formula**. If $\square = (i, j)$ is a box of the Young diagram of λ, its hook length is $h(\square) = \lambda_i + \lambda'_j - i - j + 1$. For instance, in $(5, 4, 2, 2)$, the hook length of the box $(3, 1)$ appearing on the first row and the third column is 4. Then,

$$\dim \lambda = \frac{n!}{\prod_{\square \in \lambda} h(\square)},$$

see for instance [Mac95, §1.7]. This can be proved by relating the character theory of the symmetric groups to the theory of symmetric functions that we are going to review in the next paragraph.

1.2 The Hopf algebra of symmetric functions

A formal power series[1] f of bounded degree in an infinite sequence of variables $X = \{x_1, \ldots, x_n, \ldots\}$ is called a **symmetric function** if, for any finite permutation $\sigma \in \mathfrak{S}_\infty$,

$$f(x_{\sigma(1)}, x_{\sigma(2)}, \ldots) = f(x_1, x_2, \ldots) = f(X).$$

In other words, for any $i_1 \neq i_2 \neq \cdots \neq i_r$ and $j_1 \neq j_2 \neq \cdots \neq j_r$, and any $(a_1, \ldots, a_r) \in \mathbb{N}^r$,

$$[x_{i_1}^{a_1} x_{i_2}^{a_2} \cdots x_{i_r}^{a_r}](f) = [x_{j_1}^{a_1} x_{j_2}^{a_2} \cdots x_{j_r}^{a_r}](f).$$

The set of symmetric functions will be denoted by **Sym**; this is an infinite-dimensional complex commutative algebra. A transcendence basis of **Sym** consists in the **power sums**

$$p_k(X) = \sum_{i=1}^{\infty} (x_i)^k.$$

If $\mu = (\mu_1, \ldots, \mu_r)$ is a partition, we shall denote $p_\mu = p_{\mu_1} \cdots p_{\mu_r}$; a linear basis of **Sym** is therefore $(p_\mu)_\mu$, where μ runs over $\mathfrak{Y} = \bigsqcup_{n \in \mathbb{N}} \mathfrak{Y}_n$ ([Mac95, §1.2]).

Another important basis of **Sym** is given by the **Schur functions**, see [Mac95, §1.3]. If λ is a partition, $\ell(\lambda)$ will denote its number of parts. Then, given variables $x_1, \ldots, x_N$ with $N \geq \ell(\lambda)$, we define the antisymmetric polynomial of type λ by

$$a_\lambda(x_1, x_2, \ldots, x_N) = \sum_{\sigma \in \mathfrak{S}_N} \varepsilon(\sigma) \, x^{\sigma(\lambda)},$$

---

[1] In the following, all the coefficients of the functions and power series are supposed complex, though one could define everything over $\mathbb{Q}$.

where $x^{\sigma(\lambda)} = x_1^{\lambda_{\sigma(1)}} x_2^{\lambda_{\sigma(2)}} \cdots x_N^{\lambda_{\sigma(N)}}$ — here, we complete λ with parts of size 0 to obtain a sequence of length N. Finally, the Schur polynomial is defined as the quotient

$$s_{\lambda,N}(x_1, x_2, \ldots, x_N) = \frac{a_{\lambda+\delta}(x_1, x_2, \ldots, x_N)}{a_\delta(x_1, x_2, \ldots, x_N)},$$

where $\delta = (N-1, N-2, \ldots, 1, 0)$, and

$$\lambda + \delta = (\lambda_1 + N - 1, \lambda_2 + N - 2, \ldots, \lambda_{N-1} + 1, \lambda_N).$$

If M is a matrix in $\mathrm{GL}(N, \mathbb{C})$ with eigenvalues $x_1, \ldots, x_N$, then by Weyl's character formula, $s_{\lambda,N}(x_1, \ldots, x_N)$ is the character of the action of M on the irreducible $\mathrm{GL}(N, \mathbb{C})$-module of highest weight λ. On the other hand, for any N,

$$s_{\lambda,N+1}(x_1, \ldots, x_N, x_{N+1} = 0) = s_{\lambda,N}(x_1, \ldots, x_N),$$

so there is a formal power series $s_\lambda(x_1, x_2, \ldots)$ that is the inductive limit of the $s_{\lambda,N}$'s. Since these polynomials are defined as quotients of antisymmetric polynomials, s_λ is always a symmetric function in **Sym**. Actually, $(s_\lambda)_{\lambda \in \mathfrak{Y}}$ is a linear basis of **Sym**, and by using the Schur-Weyl duality between $\mathrm{GL}(N, \mathbb{C})$ and $\mathfrak{S}_n$ on $(\mathbb{C}^N)^{\otimes n}$, one can show that the transition matrix between the two bases $(s_\lambda)_{\lambda \in \mathfrak{Y}}$ and $(p_\mu)_{\mu \in \mathfrak{Y}}$ is given by the tables of characters of the symmetric groups (see [Mac95, §1.7]).

Proposition 1. *For λ and μ in $\mathfrak{Y}_n$, we denote by $\zeta^\lambda(\mu)$ the value of the (non-normalized) character of V^λ on a permutation of cycle type μ. Then, for any partition μ,*

$$p_\mu(X) = \sum_{\lambda \in \mathfrak{Y}_n} \zeta^\lambda(\mu) \, s_\lambda(X).$$

This result is the celebrated **Frobenius formula**.

The aforementioned link between symmetric functions and the representation theory of the symmetric groups gives rise to an isomorphism between two self-adjoint **Hopf algebras**, see [Zel81]. On the one hand, let us denote by $K(\mathfrak{S}_n)$ the complex vector space built from the Grothendieck group of the category of representations of $\mathfrak{S}_n$; it means that $K(\mathfrak{S}_n)$ has for linear basis the irreducible modules V^λ with $\lambda \in \mathfrak{Y}_n$. The graded vector space $K(\mathfrak{S}) = \bigoplus_{n \in \mathbb{N}} K(\mathfrak{S}_n)$ is a graded connected Hopf algebra for the following operations:

product: $\forall \lambda \in \mathfrak{Y}_p, \ \forall \mu \in \mathfrak{Y}_q, \ V^\lambda \times V^\mu = \mathrm{Ind}_{\mathfrak{S}_p \times \mathfrak{S}_q}^{\mathfrak{S}_{p+q}}(V^\lambda \otimes V^\mu)\,;$

coproduct: $\forall \lambda \in \mathfrak{Y}_n, \ \Delta(V^\lambda) = \sum_{p+q=n} \mathrm{Res}_{\mathfrak{S}_p \times \mathfrak{S}_q}^{\mathfrak{S}_n}(V^\lambda)\,;$

involution: $\forall \lambda \in \mathfrak{Y}_n, \ \nu(V^\lambda) = V^{1^n} \otimes_\mathbb{C} V^\lambda = V^{\lambda'}.$

The involution ν defined above is equal to the antipode of the Hopf algebra up to sign. Moreover, there is a scalar product on $K(\mathfrak{S})$ defined by $\langle V^\lambda \,|\, V^\mu \rangle = \delta_{\lambda\mu}$, and $K(\mathfrak{S})$ is a self-adjoint Hopf algebra with respect to this pairing. On the other hand, **Sym** has also a structure of self-adjoint Hopf algebra if one sets:

$$\Delta(p_k) = p_k \otimes 1 + 1 \otimes p_k \qquad ; \qquad \nu(p_k) = (-1)^{k-1} p_k \qquad ; \qquad \langle p_\mu \,|\, p_\rho \rangle = \delta_{\mu\rho}\, z_\mu$$

where $z_\mu = \prod_{i \geq 1} i^{m_i}\, m_i!$ if $\mu = 1^{m_1} 2^{m_2} \cdots s^{m_s}$. Then, it is shown in [Zel81] that:

Proposition 2. *The rule $V^\lambda \mapsto s_\lambda$ provides an isomorphism of self-adjoint graded Hopf algebras between $K(\mathfrak{S})$ and* **Sym**. *With some additional conditions, this isomorphism is unique up to the multiplication by the antipode.*

In particular, the Schur functions form an orthonormal basis of **Sym**, and the character values of the symmetric groups satisfy $\zeta^\lambda(\mu) = \langle s_\lambda \,|\, p_\mu \rangle$.

In **Sym**, the structure of Hopf algebra can be related to operations on alphabets. Indeed, given two infinite alphabets X and Y, if $\overline{X} = \{-x_1, -x_2, \ldots\}$ and $X + Y = \{x_1, x_2, \ldots, y_1, y_2, \ldots\}$, then

$$(\Delta p_k)(X, Y) = p_k(X) + p_k(Y) = p_k(X + Y) \qquad ; \qquad (\nu(p_k))(X) = -p_k(\overline{X}).$$

Hence, the coproduct Δ in **Sym** corresponds simply to the addition of alphabets, and if the (formal) difference of two alphabets is defined by $p_k(X - Y) = p_k(X) - p_k(Y)$ for any $k \geq 1$, then for any symmetric function f,

$$f(X - Y) = [(1 \otimes \nu) \circ \Delta] f(X, \overline{Y}).$$

To conclude this paragraph, let us speak about **specializations**. By a formal specialization of **Sym**, we just mean a morphism of $\mathbb{C}$-algebras between **Sym** and $\mathbb{C}$. Since **Sym** $= \mathbb{C}[p_1, p_2, \ldots, p_k, \ldots]$, it amounts to give the images of the p_k's. Another way to define a specialization is to give values to the variables of the alphabet X; however, not all the specializations of **Sym** can be obtained in this way.

Example. We set

$$X_q = \frac{1}{1 - [q]} = (1, q, q^2, q^3, \ldots).$$

Then, $p_k(X_q) = \frac{1}{1 - q^k}$ for any k, and there is a hook-length formula for the specializations $s_\lambda(X_q)$ of the Schur functions corresponding to this alphabet:

$$s_\lambda(X_q) = \frac{q^{n(\lambda)}}{\prod_{\square \in \lambda} 1 - q^{h(\square)}},$$

where $n(\lambda) = \sum_{i=1}^{\ell(\lambda)} (i-1)\lambda_i$.

Given two specializations denoted by formal alphabets A and B, one can as before construct new specializations $A + B$ and $A - B$ by setting:

$$\forall k \geq 1, \ \ p_k(A + B) = p_k(A) + p_k(B) \qquad ; \qquad p_k(A - B) = p_k(A) - p_k(B).$$

Even if A and B are true alphabets, the specialization $A - B$ may not come from a true alphabet; hence, differences of alphabets give rise to "new" specializations of **Sym**, and we will see hereafter that this operation is important in the setting of the representation theory of $\mathfrak{S}_\infty$. Another important specialization of **Sym** that will be used in the next paragraph is the so-called exponential alphabet E. This formal alphabet satisfies:

$$p_1(E) = 1 \qquad ; \qquad \forall k \geq 2, \ \ p_k(E) = 0.$$

Using the reciprocal of the Frobenius formula, one sees that for any $\lambda \in \mathfrak{Y}_n$, $s_\lambda(E) = \frac{\dim \lambda}{n!}$.

1.3 Coherent systems of measures and Thoma's theorem for the characters of $\mathfrak{S}_\infty$

We denote by $\mathfrak{S}_\infty$ the set of finite permutations of $\mathbb{N}^* = [\![1, +\infty]\!]$; this is also the inductive limit of the $\mathfrak{S}_n$'s with respect to the natural embeddings $\mathfrak{S}_n \hookrightarrow \mathfrak{S}_{n+k}$. As explained in the introduction, there is no hope to have a complete knowledge of the representation theory of this group: indeed, the existence of factor representations of type II or III implies that a unitary representation of $\mathfrak{S}_\infty$ does not in general have a *unique* expansion as a direct integral of factors. However, one can classify the **characters** of $\mathfrak{S}_\infty$, that is to say the functions $\chi : \mathfrak{S}_\infty \to \mathbb{C}$ that satisfy the three following conditions:

1. χ is constant on conjugacy classes: $\forall \sigma, \tau, \ \chi(\tau \sigma \tau^{-1}) = \chi(\sigma)$.

2. χ is normalized so that $\chi(1) = 1$.

3. χ is non-negative definite: $\forall \sigma_1, \ldots, \sigma_n, \ (\chi(\sigma_i \sigma_j^{-1}))_{1 \leq i,j \leq n}$ is hermitian and non-negative definite.

Such a character is always the trace of a unitary representation of $\mathfrak{S}_\infty$ in a Von Neumann algebra of finite type. If one considers only finite type *factor* representations of $\mathfrak{S}_\infty$, then one obtained the **extremal characters** that satisfy the additional condition:

4. χ cannot be written as a combination $a\chi_1 + (1 - a)\chi_2$ with $\chi_1 \neq \chi_2$ characters and $a \in]0, 1[$.

We will denote by $\mathcal{X}(\mathfrak{S}_\infty)$ the set of characters of $\mathfrak{S}_\infty$, and by $\mathcal{X}^*(\mathfrak{S}_\infty)$ the set of extremal characters. Then, $\mathcal{X}(\mathfrak{S}_\infty)$ is a convex compact set (for the topology of simple convergence), and

$$\mathcal{X}^*(\mathfrak{S}_\infty) = \mathrm{Extr}(\mathcal{X}(\mathfrak{S}_\infty)) = \{\text{extremal points of } \mathcal{X}(\mathfrak{S}_\infty)\} \, ;$$
$$\mathcal{X}(\mathfrak{S}_\infty) = \mathscr{P}(\mathcal{X}^*(\mathfrak{S}_\infty)) = \{\text{borelian probability measures on } \mathcal{X}^*(\mathfrak{S}_\infty)\}.$$

Hence, in order to understand the character theory of $\mathfrak{S}_\infty$, it is sufficient to determine the set of extremal (or irreducible) characters of $\mathfrak{S}_\infty$. As explained beautifully in [KOV04, Oko97], this has been done by E. Thoma ([Tho64]), who relates this problem to the classification of totally positive sequences. We define the **Thoma simplex** Ω as the set of pairs $\omega = (\alpha, \beta)$ of non-increasing sequences of non-negative real numbers

$$\alpha = (\alpha_1 \geq \alpha_2 \geq \cdots \geq 0) \qquad ; \qquad \beta = (\beta_1 \geq \beta_2 \geq \cdots \geq 0)$$

such that $\sum_{i=1}^\infty \alpha_i + \sum_{i=1}^\infty \beta_i = 1 - \gamma$ with $\gamma \geq 0$. To any point ω of this infinite dimensional simplex, we associate the formal alphabet $A - B + \gamma E$, where $A = \alpha = \{\alpha_1, \alpha_2, \ldots\}$, $B = \overline{\beta} = \{-\beta_1, -\beta_2, \ldots\}$, and E is the exponential alphabet. Thus,

$$p_1(A - B + \gamma E) = p_1(\alpha) - p_1(\overline{\beta}) + \gamma p_1(E) = p_1(\alpha) + p_1(\beta) + \gamma = 1 \,;$$
$$p_{k \geq 2}(A - B + \gamma E) = p_k(\alpha) - p_k(\overline{\beta}) = p_k(\alpha) + (-1)^{k-1} p_k(\beta)$$
$$= \sum_{i=1}^\infty (\alpha_i)^k + (-1)^{k-1} \sum_{i=1}^\infty (\beta_i)^k.$$

Theorem 3. *The set of extremal characters $\mathcal{X}^*(\mathfrak{S}_\infty)$ can be identified (homeomorphically) with the Thoma simplex Ω, and the character χ^ω corresponding to a point $\omega = (\alpha, \beta) \in \Omega$ writes as*

$$\chi^\omega(\sigma) = p_\mu(A - B + \gamma E)$$

if $\sigma \in \mathfrak{S}_n$ is of cycle type $\mu \in \mathfrak{Y}_n$ — since $p_1(A - B + \gamma E) = 1$, this definition does not depend on the choice of n.

In particular, a remarkable property of extremal characters of $\mathfrak{S}_\infty$ is that they factorize on disjoint cycles. Actually, it has been proved that these two conditions are equivalent, see [Tho64, KV84, Oko97], and also [GK10] for a proof that relates this property to the notion of exchangeability of (non-commutative) random variables. Assuming only this multiplicativity, the methods of Kerov and Vershik (which are reviewed in Section 4) allow to reprove Thoma's theorem. That said, we shall assume in the following that Theorem 3 is already known, and use it as the starting point of our new results.

Another way of describing $\mathcal{X}(\mathfrak{S}_\infty)$ and $\mathcal{X}^*(\mathfrak{S}_\infty)$ is by using the notion of **coherent systems of measures** on $\mathfrak{Y} = \bigsqcup_{n \in \mathbb{N}} \mathfrak{Y}_n$. A function H on $\mathfrak{Y}$ is said harmonic if, for any partition λ,

$$H(\lambda) = \sum_{\lambda \nearrow \Lambda} H(\Lambda).$$

Let us suppose that H is positive and that $H(\emptyset) = 1$; we then set $P_n(\lambda) = H(\lambda) \times \dim \lambda$. In this setting, it is an easy recurrence to show that on each

level $\mathfrak{Y}_n$ of the Young graph $\mathfrak{Y}$, P_n is a probability measure. Indeed, for any n,

$$\sum_{\Lambda \in \mathfrak{Y}_{n+1}} P_{n+1}(\Lambda) = \sum_{\Lambda \in \mathfrak{Y}_{n+1}} H(\Lambda) \dim \Lambda = \sum_{\Lambda \in \mathfrak{Y}_{n+1}} \sum_{\lambda \nearrow \Lambda} H(\Lambda) \dim \lambda$$

$$= \sum_{\lambda \in \mathfrak{Y}_n} H(\lambda) \dim \lambda = \sum_{\lambda \in \mathfrak{Y}_n} P_n(\lambda).$$

We call coherent system of probability measures on $\mathfrak{Y}$ a sequence of probability measures $(P_n)_{n \in \mathbb{N}}$ coming from a positive normalized harmonic function on $\mathfrak{Y}$.

Proposition 4. *The coherent systems of probability measures provide a bijection between $\mathcal{X}(\mathfrak{S}_\infty)$ and the set of positive normalized harmonic functions on $\mathfrak{Y}$.*

Proof. If χ is character of $\mathfrak{S}_\infty$, its restrictions $\chi_n = \chi_{|\mathfrak{S}_n}$ are non-negative normalized characters of the symmetric groups, so they are positive linear combinations of the normalized irreducible characters χ^λ:

$$\forall n, \ \chi_{|\mathfrak{S}_n} = \sum_{\lambda \in \mathfrak{Y}_n} P_{n,\chi}(\lambda) \, \chi^\lambda.$$

Since $(\chi_{|\mathfrak{S}_{n+1}})_{|\mathfrak{S}_n} = \chi_{|\mathfrak{S}_n}$, the $P_{n,\chi}$'s form a coherent system. Conversely, given a coherent system of probability measures, one can define on each group $\mathfrak{S}_n$ a non-negative character by using the formula written above. The coherence of the system implies that these definitions are compatible, whence a well-defined character of $\mathfrak{S}_\infty$. $\qquad\square$

Let us denote by $(\mathbb{P}_{n,\omega})_{n \in \mathbb{N}}$ the coherent system associated to a point ω in the Thoma simplex. For any permutation $\sigma \in \mathfrak{S}_n \subset \mathfrak{S}_\infty$ of cycle type $\mu \in \mathfrak{Y}_n$,

$$\sum_{\lambda \in \mathfrak{Y}_n} \mathbb{P}_{n,\omega}[\lambda] \, \chi^\lambda(\mu) = \chi_{|\mathfrak{S}_n}^\omega(\sigma) = \chi^\omega(\sigma) = p_\mu(A - B + \gamma E)$$

$$=_{\text{(Frobenius formula)}} \sum_{\lambda \in \mathfrak{Y}_n} s_\lambda(A - B + \gamma E) \, \zeta^\lambda(\mu).$$

Since this is true for any permutation, the coherent system of probability measures on $\mathfrak{Y}$ associated to ω is therefore given by the specialization $A - B + \gamma E$ of the Schur functions:

Proposition 5. $\mathbb{P}_{n,\omega}[\lambda] = (\dim \lambda) \, s_\lambda(A - B + \gamma E).$

Example. Take $\omega = \mathbf{0} = ((0,0,\ldots),(0,0,\ldots))$. Then, $\chi^{\mathbf{0}}$ is the regular trace τ of the infinite symmetric group:

$$\tau(\omega) = \begin{cases} 1 & \text{if } \omega = \mathrm{id}_{\mathbb{N}}, \\ 0 & \text{otherwise.} \end{cases}$$

The corresponding measures are the so-called **Plancherel measures** of the symmetric groups:

$$\mathbb{P}_n[\lambda] = \mathbb{P}_{n,\mathbf{o}}[\lambda] = (\dim \lambda)\, s_\lambda(E) = \frac{(\dim \lambda)^2}{n!}.$$

The asymptotics of these measures have been studied in [LS77, KV77, Ker93, IO02]. It turns out that this point of the Thoma simplex will be the only one for which our methods won't give a sufficiently precise result (more on this at the end of the paper).

Proposition 5 is very important from a combinatorial point of view: it enables us to relate the coherent systems of probability measures coming from characters of $\mathfrak{S}_\infty$ to very general models of random permutations. The goal of the next section is to explain this link.

2 Quasi-symmetric functions and a link with random permutations

In this section, we shall explain how the coherent systems of measures on $\mathfrak{Y}$ are related to models of random permutations that can be defined by using generalized riffle shuffles in the sense of [Sta01, Ful02].

2.1 The Hopf algebras **QSym** and **FQSym**

To begin with, we have to present two new combinatorial Hopf algebras, namely, **QSym** and **FQSym**. In the following, $X = \{x_1, x_2, \ldots\}$ is an infinite ordered alphabet, but this time with non-commuting letters. The algebra $\mathbb{C}\langle X \rangle$ generated by these noncommutative variables has for linear basis the words $w = x_{i_1} x_{i_2} \ldots x_{i_r}$. The algebra of **free quasi-symmetric functions**, first introduced in [MR95] and redefined in [DHT01], is a subalgebra of $\mathbb{C}\langle X \rangle$ that happens to have a natural structure of Hopf algebra and to be related to numerous other combinatorial Hopf algebras. It is defined by using the **standardization** of words. Given a word $w = x_{i_1} x_{i_2} \ldots x_{i_r}$, let us denote by $\{j_1 < j_2 < \cdots < j_s\}$ the set of distinct indices i appearing in w, and by $m_1, \ldots, m_s$ their multiplicities. In w, we replace from left to right the occurrences of x_{j_1} by $1, 2, \ldots, m_1$; then, we replace from left to right the occurrences of x_{j_2} by $m_1 + 1, m_1 + 2, \ldots, m_1 + m_2$; *etc.* We thus obtain a permutation of size n, called the standardization of the word w.

Example. If $w = x_4 x_3 x_1 x_5 x_1 x_1 x_3 x_4$, then $\mathrm{Std}(w) = 64182357$.

Now, for any permutation $\sigma \in \mathfrak{S}_n$, we define elements F_σ and G_σ in $\mathbb{C}\langle X \rangle$ by

$$G_\sigma = \sum_{w \,\mid\, \mathrm{Std}(w)=\sigma} w \qquad ; \qquad F_\sigma = G_{\sigma^{-1}}.$$

It turns out that the F_σ's with $\sigma \in \mathfrak{S} = \bigsqcup_{n \in \mathbb{N}} \mathfrak{S}_n$ generate linearly a subalgebra of $\mathbb{C}\langle X \rangle$. For two permutations $\sigma = \sigma(1) \cdots \sigma(p)$ and $\tau = \tau(1) \cdots \tau(q)$, we define their shifted **shuffle product** $\sigma \,\overline{\shuffle}\, \tau$ as the set of permutations υ of size $p + q$ such that $\sigma(1), \sigma(2), \ldots, \sigma(p)$ appear in this order in the word of υ, and $\tau(1) + p, \tau(2) + p, \ldots, \tau(q) + p$ also appear in this order in the word of υ.

Example. If $\sigma = 21$ and $\tau = 231$, then

$$\sigma \,\overline{\shuffle}\, \tau = 21 \shuffle 453$$
$$= \{21453, 24153, 24513, 24531, 42153, 42513, 42531, 45213, 45231, 45321\}.$$

For all permutations $\sigma, \tau \in \mathfrak{S}$, by sorting the words appearing in the product $F_\sigma F_\tau$ according to their standardization, one sees readily that $F_\sigma F_\tau = \sum_{\upsilon \in \sigma \overline{\shuffle} \tau} F_\upsilon$.

Let us denote by **FQSym** the algebra generated by the F_σ's. The ordered sum of ordered alphabets $X + Y$ defines a coproduct on **FQSym**: $\Delta f(X, Y) = f(X + Y)$. Endowed with this new operation, **FQSym** becomes an Hopf algebra (see [DHT01, §3] and [AS05]) which is neither commutative nor cocommutative. Moreover, **FQSym** is self-dual with respect to the pairing

$$\langle F_\sigma \mid G_\tau \rangle = \delta_{\sigma\tau},$$

and in the basis $(F_\sigma)_{\sigma \in \mathfrak{S}}$, the coproduct of **FQSym** can be written in terms of deconcatenation and standardization. More precisely,

$$\Delta F_\sigma = \sum_{\sigma = \tau \cdot \nu} F_{\mathrm{Std}(\tau)} \otimes F_{\mathrm{Std}(\nu)}.$$

Example. If σ is in $\mathfrak{S}_n$, one will always obtain $n + 1$ terms in $\Delta(F_\sigma)$. For instance,

$$\Delta(F_{4132}) = 1 \otimes F_{4132} + F_1 \otimes F_{132} + F_{21} \otimes F_{21} + F_{312} \otimes F_1 + F_{4132} \otimes 1.$$

If $\sigma \in \mathfrak{S}_n$, its **descents** are the i's in $[\![1, n-1]\!]$ such that $\sigma(i) > \sigma(i+1)$, and its **recoils** are the descents of σ^{-1}. For instance, if $\sigma = 64182357$, then

$$D(\sigma) = \{1, 2, 4\} \qquad ; \qquad R(\sigma) = \{3, 5, 7\}.$$

To a subset $I \subset [\![1, n-1]\!]$, one can associate bijectively a **composition** of size n, that is to say a sequence of positive integers $c = (c_1, c_2, \ldots, c_r)$ such that $c_1 + \cdots + c_r = n$. Indeed, such a sequence is entirely determined by the integers

$$c_1 < c_1 + c_2 < \cdots < c_1 + \cdots + c_{r-1},$$

and these integers form an arbitrary subset of $[\![1, n-1]\!]$. We will denote by $c(\sigma)$ the composition associated to the descent set of a permutation σ. It is easy to compute $c(\sigma)$ by writing down the **ribbon** of σ, which is the unique ribbon tableau with reading word σ and increasing rows and columns.

Example. The ribbon of 64182357 is

$$\begin{array}{|c|c|c|c|c|}
\hline
6 \\ \cline{1-1}
4 \\ \cline{1-1}
1 & 8 \\ \hline
 & 2 & 3 & 5 & 7 \\ \hline
\end{array}.$$
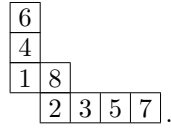
Then, the parts of $c(\sigma)$ are the lengths of the rows of the ribbon of σ. In the following, for any integer n, we will denote $\mathfrak{C}_n$ the set of compositions of size n, and if $c \in \mathfrak{C}_n$, we will denote $D(c)$ the associated subset of $[\![1, n-1]\!]$ — it is called the descent set of c.

The algebra of **quasi-symmetric functions QSym** (see for instance [Ges94] and [Sta91, §7.19]) is obtained from **FQSym** by making the variables x_i commutative. If X is now a commutative alphabet, then it is not hard to convince oneself that G_σ and G_τ become equal if and only if σ and τ have the same recoils. More precisely, under the morphism $\mathbb{C}\langle X \rangle \to \mathbb{C}[X]$, a free quasi-symmetric function F_σ becomes $L_{c(\sigma)}$, where for any composition $c \in \mathfrak{C}_n$, L_c is defined as

$$L_c = \sum_{\substack{i_1 \leq i_2 \leq \cdots \leq i_n \\ i_j < i_{j+1} \text{ if } j \in D(c)}} x_{i_1} x_{i_2} \cdots x_{i_n}.$$

A formal power series f in commutative variables $x_1, x_2, \ldots$ belongs to the vector space generated by the L_c's, $c \in \mathfrak{C} = \bigsqcup_{n \in \mathbb{N}} \mathfrak{C}_n$ if and only if it is quasi-symmetric, meaning that for any $i_1 < i_2 < \cdots < i_r$ and $j_1 < j_2 < \cdots < j_r$, and any $(a_1, \ldots, a_r) \in \mathbb{N}^r$,

$$[x_{i_1}^{a_1} x_{i_2}^{a_2} \cdots x_{i_r}^{a_r}](f) = [x_{j_1}^{a_1} x_{j_2}^{a_2} \cdots x_{j_r}^{a_r}](f).$$

The reader should compare this definition of quasi-symmetric functions with the definition of symmetric functions given on page 6. The space **QSym** of quasi-symmetric functions is stable by product, and in fact, it is again an Hopf algebra; it contains **Sym** as a Hopf subalgebra, and it is the quotient of **FQSym** by the congruences

$$F_\sigma \equiv F_\tau \quad \Longleftrightarrow \quad c(\sigma) = c(\tau).$$

The product of two fundamental quasi-symmetric functions L_c and L_d is not in general multiplicity-free as in **FQSym**[2]; however, the formula for a coproduct $\Delta(L_c)$ remains essentially the same as the one given in **FQSym** for $\Delta(F_\sigma)$. More precisely, for $c \in \mathfrak{C}_n$ and $i, j \in [\![0, n]\!]$, let us denote $c_{[\![i+1,j]\!]}$ the composition of size $j - i$ that satisfies:

$$D(c_{[\![i+1,j]\!]}) = \big(D(c) \cap [\![i+1, j-1]\!] \big) - i.$$

In terms of ribbons, it amounts to cut the ribbon of c before the box $i+1$ and after the box j.

---

[2] This is one the main advantage of doing computations in **FQSym**.

Example. If $c = (1, 1, 2, 4)$ and $[\![i+1, j]\!] = [\![3, 7]\!]$, then the ribbon of c is
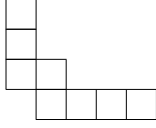
and it is then easy to see that $c_{[\![3,7]\!]} = (2, 3)$.

The translation in **QSym** of the formula given previously for a coproduct $\Delta(F_\sigma)$ is quite obviously:

$$\Delta(L_c) = \sum_{i=0}^{n} L_{c_{[\![1,i]\!]}} \otimes L_{c_{[\![i+1,n]\!]}}.$$

Example.

$$\Delta(L_{1,2,1}) = L_\emptyset \otimes L_{1,2,1} + L_1 \otimes L_{2,1} + L_{1,1} \otimes L_{1,1} + L_{1,2} \otimes L_1 + L_{1,2,1} \otimes L_\emptyset.$$

We shall also need the formula for the antipode of **QSym**. If $c = (c_1, \ldots, c_r)$ is a composition, let us denote by c^* the composition whose descent set is $(c_r, c_r + c_{r-1}, \ldots, c_r + \cdots + c_2)$; this is the reverse composition of c. Then, we define the conjugate $\bar{c}$ of c by $D(\bar{c}) = [\![1, n-1]\!] \setminus D(c^*)$. The operation $c \mapsto \bar{c}$ may seem complicate, but in terms of ribbons, it amounts simply to symmetrize c with respect to the first diagonal.

Example. The conjugate of $c = (1, 1, 2, 4)$ has for ribbon
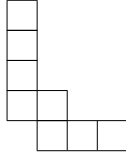
so $\bar{c} = (1, 1, 1, 2, 3)$.

Then, it can be shown that the antipode of **QSym** is up to sign equal to ν, where $\nu(L_c) = L_{\bar{c}}$ for any composition. In **FQSym**, the antipode of is much more difficult to compute; see [AS05, §5]. Fortunately, we shall only need the following property:

$$\nu(F_{12\ldots n}) = F_{n\ldots 21}$$

which is an immediate consequence[3] of [AS05, Theorem 5.4].

To conclude this paragraph, let us explain how the bases of **Sym**, **QSym** and **FQSym** are related. If $\sigma \in \mathfrak{S}_n$ is a permutation, the Robinson-Schensted-Knuth (RSK) algorithm ([Rob38, Sch61]) provides a pair of standard tableaux

---

[3] This partial rule is compatible with the following identity in **QSym**: for any permutation $\sigma \in \mathfrak{S}_n$, $\nu(L_{c(\sigma)}) = L_{c(\sigma\omega_0)}$, where ω_0 is the permutation of maximal length in $\mathfrak{S}_n$.

$(P(\sigma), Q(\sigma))$ of same shape $\Lambda(\sigma) \in \mathfrak{Y}_n$, and it is actually a bijection between $\mathfrak{S}_n$ and $\bigsqcup_{\lambda \in \mathfrak{Y}_n} \mathrm{ST}(\lambda)^2$. Given a permutation[4] $\sigma = w_1 w_2 \cdots w_n$, one constructs the first tableau $P(\sigma)$ as follows:

1. One reads the word from left to right, and one inserts the letters successively in the standard tableau P, starting from $P = \emptyset$.

2. When inserting the letter a, one compare a to the entries of the first row. If a is greater than all the entries of the first row, one adds a at the end of the row. Otherwise, if b is the first entry strictly greater than a, one replaces b by a in the first row, and one inserts b in the next row (with the same rules).

The tableau $Q(\sigma)$ is constructed by recording the order of creation of the boxes of P. The correspondence $\sigma \mapsto (P(\sigma), Q(\sigma))$ has the following additional properties ([Ful97]):

1. For any permutation σ, $P(\sigma) = Q(\sigma^{-1})$.

2. The parts of the common shape $\Lambda(\sigma)$ of $P(\sigma)$ and $Q(\sigma)$ are related to the so-called **Greene invariants** of the permutation (see *e.g.* [LLT02]). Hence, for any $i \leq \ell(\Lambda(\sigma))$,

$$\Lambda_1 + \Lambda_2 + \cdots + \Lambda_i = \max\{\ell(iw_1) + \ell(iw_2) + \cdots + \ell(iw_i)\}$$

where the maximum is taken on the set of i-tuples of disjoint increasing subwords of σ. In particular, Λ_1 is the length $\ell(\sigma)$ of a longest increasing subword of σ. One has the same result if one replaces $\Lambda(\sigma)$ by its conjugate and the longest increasing subwords by the longest decreasing subwords.

3. A descent of a standard tableau T is an index i such that $i+1$ appears in a row strictly above the one that contains i. This notion allows to associate to T a subset of $[\![1, n-1]\!]$, and therefore a composition $c(T)$. Then, for any permutation σ, $c(\sigma) = c(Q(\sigma))$.

Example. If $\sigma = 64182357$, the two corresponding tableaux are

$$
\begin{array}{l}
\boxed{6} \\
\boxed{4}\,\boxed{8} \\
\boxed{1}\,\boxed{2}\,\boxed{3}\,\boxed{5}\,\boxed{7}
\end{array}
\quad \text{and} \quad
\begin{array}{l}
\boxed{3} \\
\boxed{2}\,\boxed{5} \\
\boxed{1}\,\boxed{4}\,\boxed{6}\,\boxed{7}\,\boxed{8}
\end{array}.
$$

The length of the first part is 5, and it is indeed the length of a longest increasing subword, here, 12357. The descents of $Q(\sigma)$ are $\{1, 2, 4\}$, and they are indeed the descents of σ.

---

[4] Or more generally, a word w; then, $P(w)$ is a *semi-standard* tableau whose entries are the letters of w.

Let us denote $\pi : \mathbf{FQSym} \to \mathbf{QSym}$ the quotient map coming from the morphism $\mathbb{C}\langle X \rangle \to \mathbb{C}[X]$. For any n, notice that in $\mathbf{FQSym}$,

$$(F_1)^n = \sum_{\sigma \in \mathfrak{S}_n} F_\sigma.$$

We decompose this sum in blocks according to the shape $\Lambda(\sigma)$ of the two standard tableaux corresponding to σ:

$$(F_1)^n = \sum_{\lambda \in \mathfrak{Y}_n} \left(\sum_{\sigma \ | \ \Lambda(\sigma) = \lambda} F_\sigma \right).$$

If one applies π to the left-hand side of this expansion, one obtains $(L_{(1)})^n$. On the other hand, if one applies π to a block B_λ in the right hand side, one gets

$$\pi(B_\lambda) = \sum_{P, Q \in \mathrm{ST}(\lambda)} L_{c(Q)} = (\dim \lambda) \sum_{Q \in \mathrm{ST}(\lambda)} L_{c(Q)}.$$

But in fact, after applying π, both members of the equation are symmetric functions (not just *quasi-symmetric*), and the decomposition is simply the Frobenius formula for $(p_1)^n$:

$$(p_1)^n = \sum_{\lambda \in \mathfrak{Y}_n} (\dim \lambda) \, s_\lambda.$$

In particular, $s_\lambda = \sum_{Q \in \mathrm{ST}(\lambda)} L_{c(Q)}$, see [Sta91, p. 361]. This simple fact allows to lift the coherent systems of measures to systems of probability measures on the symmetric groups $\mathfrak{S}_n$.

2.2 Coherent systems as push-forwards of measures on the symmetric groups

To do this, we will lift the specializations $A - B + \gamma E$ to specializations of $\mathbf{FQSym}$. First, remark that for any symmetric function f, one can write:

$$f(A - B + \gamma E) = \Delta f(A - B, \gamma E) = [(\mathrm{id} \otimes \nu \otimes \mathrm{id}) \circ (\Delta \otimes \mathrm{id}) \circ \Delta] f(A, \overline{B}, \gamma E).$$

Suppose that the specializations A, $\overline{B}$ and γE are lifted up to $\mathbf{QSym}$; this means that one has morphisms of algebras $A : \mathbf{QSym} \to \mathbb{C}$, $B : \mathbf{QSym} \to \mathbb{C}$ and $\gamma E : \mathbf{QSym} \to \mathbb{C}$, whose restrictions to $\mathbf{Sym} \subset \mathbf{QSym}$ are those defined in Theorem 3. Then, since ν and Δ are morphisms of the commutative algebra $\mathbf{QSym}$, the formula above defines a specialization of $\mathbf{QSym}$ that is compatible with the specialization $A - B + \gamma E$ of $\mathbf{Sym}$. Moreover, by composition by the quotient map $\pi : \mathbf{FQSym} \to \mathbf{QSym}$, one gets a specialization of $\mathbf{FQSym}$. Hence, it is sufficient to define $f(A)$, $f(\overline{B})$ and $f(\gamma E)$ for any quasi-symmetric function f.

For the two first, exactly as in the case of $\mathbf{Sym}$, it amounts to replace the variables $X = \{x_1, x_2, \ldots\}$ by either $A = \{\alpha_1, \alpha_2, \ldots, \}$ or $\overline{B} = \{\beta_1, \beta_2, \ldots\}$.

Notice that since $\sum_{i=1}^{\infty} \alpha_i + \sum_{i=1}^{\infty} \beta_i \leq 1$, for any $f \in \mathbf{QSym}$, the series $f(A)$ and $f(\overline{B})$ converge absolutely. As for the exponential alphabet, in $\mathbf{FQSym}$ or $\mathbf{QSym}$, it is defined by the rule:

$$F_\sigma(\gamma E) = L_{c(\sigma)}(\gamma E) = \frac{\gamma^n}{n!}.$$

Since a product $F_\sigma F_\tau$ involves $\binom{|\sigma|+|\tau|}{|\sigma|,|\tau|}$ terms, this rule gives indeed a morphism of algebras. So, the specialization $A - B + \gamma E$ is well-defined on $\mathbf{QSym}$ and $\mathbf{FQSym}$. Moreover, notice that for any permutation σ, $F_\sigma(A)$, $F_\sigma(\overline{B})$ and $F_\sigma(\gamma E)$ are non-negative. Since Δ and ν are positive operators on $\mathbf{QSym}$ with respect to the basis $(L_c)_{c \in \mathfrak{C}}$,

$$\forall \omega = (\alpha, \beta) \in \Omega, \ \ \forall \sigma \in \mathfrak{S}, \ \ F_\sigma(A - B + \gamma E) > 0.$$

On the other hand, since $F_1(A - B + \gamma E) = p_1(A - B + \gamma E) = 1$,

$$\sum_{\sigma \in \mathfrak{S}_n} G_\sigma(A - B + \gamma E) = \sum_{\sigma \in \mathfrak{S}_n} F_\sigma(A - B + \gamma E) = (F_1(A - B + \gamma E))^n = 1.$$

Thus, by using the Hopf algebras $\mathbf{QSym}$ and $\mathbf{FQSym}$, for any point ω of the Thoma simplex, one can define a (positive) probability measure $\mathbb{Q}_{n,\omega}$ on the symmetric group $\mathfrak{S}_n$ corresponding *via* RSK to the coherent measure $\mathbb{P}_{n,\omega}$:

Proposition 6. *For any $\omega \in \Omega$, and any permutation $\sigma \in \mathfrak{S}$, set $\mathbb{Q}_{n,\omega}[\sigma] = G_\sigma(A - B + \gamma E)$.*

1. For any n, $\mathbb{Q}_{n,\omega}$ is a probability measure on $\mathfrak{S}_n$, and for any partition λ,

$$\mathbb{P}_{n,\omega}[\lambda] = \Lambda_* \mathbb{Q}_{n,\omega}[\lambda] = \sum_{\sigma \,|\, \Lambda(\sigma) = \lambda} \mathbb{Q}_{n,\omega}[\sigma].$$

2. If c is the composition of size n encoding the recoils of σ, then

$$\mathbb{Q}_{n,\omega}[\sigma] = \sum_{i+j+k=n} \frac{\gamma^k}{k!} \, L_{c_{[\![1,i]\!]}}(\alpha_1, \alpha_2, \ldots) \, L_{\overline{c_{[\![i+1,i+j]\!]}}}(\beta_1, \beta_2, \ldots).$$

Proof. We have seen just before that $\mathbb{Q}_{n,\omega}[\sigma] = G_\sigma(A - B + \gamma E)$ defines a probability measure on $\mathfrak{S}_n$. Now, for any partition $\lambda \in \mathfrak{Y}_n$, we can write:

$$\sum_{\Lambda(\sigma) = \lambda} \mathbb{Q}_{n,\omega}[\sigma] = \sum_{\Lambda(\sigma) = \lambda} L_{c(\sigma^{-1})}(A - B + \gamma E) = \sum_{P,Q \in \mathrm{ST}(\lambda)} L_{c(P)}(A - B + \gamma E)$$

$$= (\dim \lambda) \, s_\lambda(A - B + \gamma E) = \mathbb{P}_{n,\omega}[\lambda].$$

Finally, the specialization $A - B + \gamma E$ of **FQSym** factors through the quotient map $\pi : \mathbf{FQSym} \to \mathbf{QSym}$, and in this latter algebra, it is not difficult to compute it:

$$
\begin{aligned}
\mathbb{Q}_{n,\omega}[\sigma] &= G_\sigma(A - B + \gamma E) = L_{c(\sigma^{-1})}(A - B + \gamma E) \\
&= [(\mathrm{id} \otimes \nu \otimes \mathrm{id}) \circ (\Delta \otimes \mathrm{id}) \circ \Delta] L_c(A, \overline{B}, \gamma E) \\
&= \sum_{i+j+k=n} L_{c_{\llbracket 1,i \rrbracket}}(A)\,(\nu L_{c_{\llbracket i+1,i+j \rrbracket}})(\overline{B})\,L_{c_{\llbracket i+j+1,n \rrbracket}}(\gamma E) \\
&= \sum_{i+j+k=n} L_{c_{\llbracket 1,i \rrbracket}}(\alpha)\,L_{\overline{c_{\llbracket i+1,i+j \rrbracket}}}(\beta)\,\frac{\gamma^k}{k!}.
\end{aligned}
$$

$\square$

Example. If $\omega = ((0,0,\ldots),(0,0,\ldots))$, then $\gamma = 1$ and $\mathbb{Q}_{n,\omega}[\sigma] = \frac{1}{n!}$ for any permutation $\sigma \in \mathfrak{S}_n$ (in the triple sum, the only non-zero term corresponds to $i = j = 0$ and $k = n$). Thus, one obtains the uniform measures.

Similarly, suppose that $\omega = ((1 - \gamma, 0, 0, \ldots),(0,0,\ldots))$ with $\gamma > 0$. Then, $L_{c_{\llbracket 1,i \rrbracket}}(\alpha)$ equals 0 unless $c_{\llbracket 1,i \rrbracket}$ has no descents, which is equivalent to ask that the first recoil of σ is after i. So, if $\mathrm{FR}(\sigma)$ denotes this first recoil, then

$$
\mathbb{Q}_{n,\omega}[\sigma] = \sum_{i=0}^{\mathrm{FR}(\sigma)} \frac{(1 - \gamma)^i\,\gamma^{n-i}}{n - i!}.
$$

Suppose now that $\omega = ((1/d, 1/d, \ldots, 1/d, 0, 0, \ldots),(0,0,\ldots))$, with d terms $1/d$ in α. Then,

$$
\mathbb{Q}_{n,\omega}[\sigma] = L_{c(\sigma)^{-1}}(\alpha) = \sum_{\substack{i_1 \leq i_2 \leq \ldots \leq i_n \leq d \\ i_j < i_{j+1} \text{ if } j \in D(c(\sigma^{-1}))}} \frac{1}{d^n} = \frac{1}{d^n} \binom{n - 1 + d - \mathrm{card}\, R(\sigma)}{n}.
$$

For the two last examples, exchanging α and β amounts to replace σ by $\sigma\omega_0$, where $\omega_0 = (n, n - 1, \ldots, 1)$.

Finally, suppose that $\omega = (((1 - q),(1 - q)q,(1 - q)q^2,\ldots),(0,0,\ldots))$. Then,

$$
\mathbb{Q}_{n,\omega}[\sigma] = L_{c(\sigma)^{-1}}(\alpha) = (1 - q)^n \sum_{\substack{i_1 \leq \ldots \leq i_n \leq d \\ i_j < i_{j+1} \text{ if } j \in D(c(\sigma^{-1}))}} q^{i_1 + \cdots + i_n - n} = \frac{q^{\mathrm{comaj}(\sigma^{-1})}}{\{n!\}_q},
$$

where $\mathrm{comaj}(\sigma) = \sum_{d \in D(\sigma)} n - d$ and $\{n!\}_q = \prod_{i=1}^{n} \frac{1 - q^i}{1 - q}$.

2.3 Link with the generalized riffle shuffles

A remarkable fact is that the probability measures $\mathbb{Q}_{n,\omega}$ defined in the previous paragraph can be described practically by random **generalized riffle shuffles**; see [Ful02]. We fix a point $\omega = (\alpha, \beta)$ of the Thoma simplex and an integer n, and we consider a deck of cards ordered from 1 to n. We then construct a random permutation $\sigma \in \mathfrak{S}_n$ by using the following algorithm:

1. First, we split the deck in blocks of size $n_1 + n_2 + \cdots + m_1 + m_2 + \cdots + l = n$, the sizes of the blocks being chosen randomly according to a multinomial law of parameter (α, β, γ):

$$\mathbb{P}[n_1, \ldots, m_1, \ldots, l] = \binom{n}{n_1, \ldots, m_1, \ldots, l} \left(\prod_{i \geq 1} \alpha_i^{n_i} \right) \left(\prod_{j \geq 1} \beta_j^{m_j} \right) \gamma^l.$$

2. In the blocks of sizes $m_1, m_2, \ldots$, we reverse the order of the cards. In the block of size l, we randomize the order of the cards, so that every permutation of the l last cards becomes equiprobable.

3. Finally, we randomly shuffle the blocks with the following rule. If two blocks of size k_1 and k_2 are shuffled together, then the first card of the new deck comes from the top of the first block with probability $\frac{k_1}{k_1+k_2}$, and from the top of the second block with probability $\frac{k_2}{k_1+k_2}$; then we start again with the remaining $k_1 + k_2 - 1$ cards. Notice that this recursive rule is equivalent to ask that all the $\binom{k_1+k_2}{k_1,k_2}$ shuffles of the blocks are equiprobable.

This algorithm is called in [Ful02] an (α, β, γ)-shuffle.

Proposition 7. *The probability of a permutation σ obtained by the ω-shuffle presented above is exactly $\mathbb{Q}_{n,\omega}[\sigma]$.*

Proof. Let us give a purely algebraic proof. If X is a formal alphabet, we denote by $\phi_X : \mathbf{FQSym} \to \mathbb{C}$ the corresponding specialization, and by u_X and t_X the maps

$$u_X : \mathbf{FQSym} \to \mathbf{FQSym} \qquad ; \qquad t_X : \mathbf{FQSym} \to \mathbf{FQSym}$$
$$F_\sigma \mapsto F_\sigma(X) \, F_{12\ldots n} \qquad\qquad\qquad F_\sigma \mapsto F_\sigma(X) \, F_\sigma.$$

Notice that if $X = (\zeta)$ is a single variable, then $u_{(\zeta)} = t_{(\zeta)}$. We start from the identity $\mathbb{Q}_{n,\omega}[\sigma] = G_\sigma(A - B + \gamma E)$, and we expand it:

$$\mathbb{Q}_{n,\omega}[\sigma] = \phi_{A-B+\gamma E}(G_\sigma) = m^3 \circ (\phi_A \otimes \phi_{-B} \otimes \phi_{\gamma E}) \circ \Delta^3(G_\sigma)$$
$$= m^3 \circ (\phi_A \otimes \phi_{\overline{B}} \otimes \phi_{\gamma E}) \circ (1 \otimes \nu \otimes 1) \circ \Delta^3(G_\sigma)$$

The operator applied to G_σ can be further expanded in

$$m^{2\infty+1} \circ \left(\left(\otimes_i \phi_{(\alpha_i)} \right) \otimes \left(\otimes_j \phi_{(\beta_j)} \right) \otimes \phi_{\gamma E} \right) \circ \left(1^{\otimes \infty} \otimes \nu^{\otimes \infty} \otimes 1 \right) \circ \Delta^{2\infty+1}$$

where the meaning of the symbols $2\infty + 1$ is quite clear from what precedes. Let us denote $I^{2\infty+1} = 1^{\otimes\infty} \otimes \nu^{\otimes\infty} \otimes 1$. Then, $\mathbb{Q}_{n,\omega}[\sigma]$ can be rewritten as:

$$\left\langle \Delta^{2\infty+1}(F_{12\ldots n}) \,\middle|\, \left(\left(\otimes_i u_{(\alpha_i)}\right) \otimes \left(\otimes_j u_{(\beta_j)}\right) \otimes u_{\gamma E} \right) \circ I^{2\infty+1} \circ \Delta^{2\infty+1}(G_\sigma) \right\rangle$$
$$= \left\langle \Delta^{2\infty+1}(F_{12\ldots n}) \,\middle|\, \left(\left(\otimes_i t_{(\alpha_i)}\right) \otimes \left(\otimes_j t_{(\beta_j)}\right) \otimes u_{\gamma E} \right) \circ I^{2\infty+1} \circ \Delta^{2\infty+1}(G_\sigma) \right\rangle$$
$$= \left\langle m^{2\infty+1} \circ I^{2\infty+1} \circ \left(\left(\otimes_i t^*_{(\alpha_i)}\right) \otimes \left(\otimes_j t^*_{(\beta_j)}\right) \otimes u^*_{\gamma E} \right) \circ \Delta^{2\infty+1}(F_{12\ldots n}) \,\middle|\, G_\sigma \right\rangle$$

Here, we have used extensively the property of self-adjointness of **FQSym**. Now, we have to check that the operator applied to $F_{12\ldots n}$ encodes the random steps of the ω-shuffle. We describe a probability distribution $\mathbb{P}$ on $\mathfrak{S}_n$ by an element $\sum_{\sigma \in \mathfrak{S}_n} \mathbb{P}[\sigma]\, F_\sigma \in$ **FQSym**, and the effect of our random steps by matrices

$$M = \sum_{\sigma, \tau} c_{\sigma\tau}\, F_\sigma \otimes G_\tau,$$

where the G's are considered as elements of the dual space of **FQSym**. We also introduce the operator of uniformization

$$U = \sum_{n \in \mathbb{N}} \sum_{\sigma \in \mathfrak{S}_n} \frac{1}{n!} F_\sigma \otimes G_{12\ldots n}.$$

Notice that for any n,

$$u^*_{\gamma E}(F_{12\ldots n}) = \frac{\gamma^n}{n!} F_{12\ldots n} = (U \circ t_{(\gamma)})(F_{12\ldots n}) \qquad ; \qquad F_{12\ldots n} = G_{12\ldots n}.$$

For this reason, we can rewrite the probability $\mathbb{Q}_{n,\omega}$ as

$$\left\langle m^{2\infty+1} \circ \widetilde{I}^{2\infty+1} \circ \left(\left(\otimes_i t_{(\alpha_i)}\right) \otimes \left(\otimes_j t_{(\beta_j)}\right) \otimes t_{(\gamma)} \right) \circ \Delta^{2\infty+1}(F_{12\ldots n}) \,\middle|\, G_\sigma \right\rangle,$$

where $\widetilde{I}^{2\infty+1} = 1^{\otimes\infty} \otimes \nu^{\otimes\infty} \otimes U$.

1. Starting from the state $F_{12\ldots n}$, the first step of our algorithm can be written

$$S_1 = \left(\left(\otimes_i t_{(\alpha_i)}\right) \otimes \left(\otimes_j t_{(\beta_j)}\right) \otimes t_{(\gamma)} \right) \circ t_E^{\otimes 2\infty+1} \circ \Delta^{2\infty+1} \circ t_{E^{-1}}.$$

Indeed, $\Delta^{2\infty+1}$ enables us to cut the deck in every possible way

$$n = n_1 + n_2 + \cdots + m_1 + m_2 + \cdots + l,$$

and the other operators take account of the probabilities (we use the $t_{(\varsigma)}$'s for the powers and the t_E's for the factorials).

2. The second step is obviously $S_2 = \widetilde{I}^{2\infty+1} = 1^{\otimes\infty} \otimes \nu^{\otimes\infty} \otimes U$. Here, we use the aforementioned fact that the involution ν reverses the term $F_{12\ldots n}$ in $F_{n(n-1)\ldots 1}$.

3. Finally, since all the shuffles are equiprobable, by the very definition of the multiplication in **FQSym**, the third step amounts to apply

$$S_3 = t_E \circ m^{2\infty+1} \circ t_{E^{-1}}^{\otimes 2\infty+1}.$$

Finally, in $S_3 \circ S_2 \circ S_1$, an operator t_E commutes with every degree-preserving operator, so one can simplify them. One gets therefore

$$S_3 \circ S_2 \circ S_1 = m^{2\infty+1} \circ \widetilde{I}^{2\infty+1} \circ \left(\left(\otimes_i t_{(\alpha_i)} \right) \otimes \left(\otimes_j t_{(\beta_j)} \right) \otimes t_{(\gamma)} \right) \circ \Delta^{2\infty+1}$$

and the proof is complete.   □

Hence, any point ω of the Thoma simplex Ω corresponds to:

1. an irreducible character χ^ω of $\mathfrak{S}_\infty$ and a coherent system of probability measures $(\mathbb{P}_{n,\omega})_{n \in \mathbb{N}}$ on the levels $\mathfrak{Y}_n$ of the Young graph;

2. a system of measures $(\mathbb{Q}_{n,\omega})_{n \in \mathbb{N}}$ on the symmetric groups $\mathfrak{S}_n$ such that $\Lambda_* \mathbb{Q}_{n,\omega} = \mathbb{P}_{n,\omega}$;

3. and a random algorithm of shuffling that produces permutations $\sigma \in \mathfrak{S}_n$ with distribution $\mathbb{Q}_{n,\omega}$.

It is now time to introduce the tools that shall enable us to perform the asymptotic analysis of these objects.

3 Observables of Young diagrams and their cumulants

For the asymptotic analysis of Young diagrams, we are basically going to use a method of moments, but we have to define precisely what we mean by the moments of a partition. We shall essentially follow [IO02], and [Śni06b, FM10] for the combinatorics of cumulants.
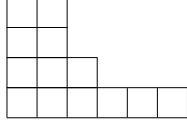
3.1 Polynomial functions on Young diagrams

If λ is a partition, we associate to it two finite alphabets $A(\lambda)$ and $B(\lambda)$ called the **Frobenius coordinates**:

$$A(\lambda) = \left\{ \lambda_1 - \frac{1}{2}, \lambda_2 - \frac{3}{2}, \ldots, \lambda_d - \frac{2d-1}{2} \right\};$$

$$B(\lambda) = A(\lambda') = \left\{ \lambda_1' - \frac{1}{2}, \lambda_2' - \frac{3}{2}, \ldots, \lambda_d' - \frac{2d-1}{2} \right\}.$$

Here, d is the size of the diagonal of λ, that is to say the number of boxes of the Young diagram of λ that belong to the first diagonal.

Example. If $\lambda = (6, 3, 2, 2)$, its Young diagram is

and the Frobenius coordinates are $A = \{\frac{11}{2}, \frac{3}{2}\}$ and $B = \{\frac{7}{2}, \frac{5}{2}\}$ — they are easily computed by drawing the diagonal and measuring the rows and columns starting from the diagonal.

The **Frobenius moments** of λ are the power sums

$$p_k(\lambda) = p_k\left(A(\lambda) - \overline{B}(\lambda)\right) = \left(\sum_{i=1}^{d} a_i^k\right) + (-1)^{k-1}\left(\sum_{i=1}^{d} b_i^k\right).$$

For instance, $p_1(\lambda) = |\lambda|$ for any partition. The functions p_k generate freely over $\mathbb{C}$ an algebra called the algebra of **observables of diagrams**, and denoted $\mathscr{O}$ (*cf.* [KO94,IO02]). This algebra is isomorphic to **Sym**, and it is naturally graded by $\deg p_k = k$. It does not contain all the polynomial functions of the coordinates of the partitions, but it will prove sufficiently large[5] in our setting.

We now define new functions on Young diagrams by renormalizing the values of the irreducible characters of the symmetric groups. Recall that for any partitions λ, μ of the same size, $\chi^\lambda(\mu)$ denotes the value of the *normalized* irreducible character of $\mathfrak{S}_n$ of label λ on a permutation of cycle type μ; thus, $\zeta^\lambda(\mu) = (\dim \lambda)\, \chi^\lambda(\mu)$. Then, for any partitions λ, μ, we set:

$$\Sigma_\mu(\lambda) = \begin{cases} n(n-1)\ldots(n-|\mu|+1)\,\chi^\lambda(\mu \sqcup 1^{n-|\mu|}) & \text{if } n = |\lambda| \geq |\mu|, \\ 0 & \text{otherwise.} \end{cases}$$

For convenience, we will denote $n(n-1)\cdots(n-k+1) = n^{\downarrow k}$ for any n and k — this is the Pochhammer symbol for falling factorials. Thus, $\Sigma_\mu(\lambda) = |\lambda|^{\downarrow|\mu|}\,\chi^\lambda(\mu \sqcup 1^{|\lambda|-|\mu|})$. An essential result exposed in details in [IO02] is the following:

Theorem 8. *Viewed as a function on Young diagrams, Σ_μ is an element of $\mathscr{O}$, and is of degree $|\mu|$. Moreover,*

$$\forall \mu, \quad \Sigma_\mu = p_\mu + (\text{terms of degree at most } |\mu| - 1),$$

see [IO02, Proposition 3.4].

---

[5] However, we should mention that there is a bigger algebra of observables of diagrams that contains $\mathscr{O}$, is isomorphic to **QSym** and allows to decompose the observables of $\mathscr{O}$ in combinatorial summands; see in particular [DFŚ10].

Example. Actually, there is an explicit formula for the Σ_k's that is due to Wassermann ([Was81], [IO02, Proposition 3.3]). Hence, one can compute $\Sigma_1 = p_1$, $\Sigma_2 = p_2$, $\Sigma_3 = p_3 - \frac{3}{2}p_{11} + \frac{5}{4}p_1$, $\Sigma_4 = p_4 - 4p_{21} + \frac{11}{2}p_2$.

Theorem 8 implies in particular that the symbols Σ_μ factorize in higher degree, that is to say that for any partitions μ and ν,

$$\Sigma_\mu \, \Sigma_\nu = \Sigma_{\mu \sqcup \nu} + \text{(terms of degree at most } |\mu| + |\nu| - 1).$$

The next paragraph is devoted to a better description of the products $\Sigma_\mu \, \Sigma_\nu$; we shall rely on the formalism of partial permutations (*cf.* [IK99]). To conclude this paragraph, let us compute the expectation of a symbol $\Sigma_\mu(\lambda)$ when λ is chosen randomly according to a measure $\mathbb{P}_{n,\omega}$ (we suppose that $n \geq |\mu|$):

$$\begin{aligned}
\mathbb{E}_{n,\omega}[\Sigma_\mu(\lambda)] &= n^{\downarrow|\mu|} \sum_{\lambda \in \mathfrak{Y}_n} \mathbb{P}_{n,\omega}[\lambda] \, \chi^\lambda(\mu \sqcup 1^{n-|\mu|}) \\
&= n^{\downarrow|\mu|} \sum_{\lambda \in \mathfrak{Y}_n} s_\lambda(A - B + \gamma E) \, \zeta^\lambda(\mu \sqcup 1^{n-|\mu|}) \\
&=_{\text{(Frobenius formula)}} n^{\downarrow|\mu|} \, p_{\mu \sqcup 1^{n-|\mu|}}(A - B + \gamma E) \\
&= n^{\downarrow|\mu|} \, p_\mu(A - B + \gamma E).
\end{aligned}$$

Theorem 8 implies that $(\Sigma_\mu)_{\mu \in \mathfrak{Y}}$ is a linear basis of $\mathscr{O}$; hence, by using this simple formula, we will be able to compute the expectation of *any* observable of diagrams under a coherent measure $\mathbb{P}_{n,\omega}$. In particular, we shall compute in Sections 4 and 5 the joint moments of the Frobenius observables p_k; they will provide information on the geometry of the random partitions under the probability measures $\mathbb{P}_{n,\omega}$.

3.2 Ivanov-Kerov algebra of partial permutations

If $\sigma \in \mathfrak{S}_\infty$, we call admissible set for σ a finite subset $S \subset \mathbb{N}^*$ such that if $x \notin S$, then $\sigma(x) = x$. A **partial permutation** is a pair (σ, S) with S admissible set for σ. Said in other words, a partial permutation is a subset $S \subset \mathbb{N}^*$ and a permutation $\sigma \in \mathfrak{S}(S)$. One can multiply these objects by the rule

$$(\sigma, S) \times (\tau, T) = (\sigma\tau, S \cup T).$$

We denote $\deg(\sigma, S) = \operatorname{card} S$, and we construct a graded algebra $\mathscr{A}_\infty$ by looking at all the formal linear combinations of partial permutations with bounded degree. For any n, this algebra $\mathscr{A}_\infty$ can be projected onto the group algebra $\mathbb{C}\mathfrak{S}_n$:

$$\begin{aligned}
\pi_n : \mathscr{A}_\infty &\to \mathbb{C}\mathfrak{S}_n \\
(\sigma, S) &\mapsto \begin{cases} \sigma & \text{if } S \subset [\![1, n]\!], \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}$$

There is a natural action of $\mathfrak{S}_\infty$ on $\mathscr{A}_\infty$: on the basis elements (σ, S), it is defined by

$$(\sigma, S)^\tau = (\tau\sigma\tau^{-1}, \tau(S)),$$

and it is compatible with the product of partial permutations. Moreover, two partial permutations (σ, S) and (τ, T) are conjugated if and only if $\operatorname{card} S = \operatorname{card} T$, and the cyclic type of $\sigma_{|S}$ is the same as the cyclic type of $\tau_{|T}$. Thus, the invariant algebra $\mathscr{B}_\infty = (\mathscr{A}_\infty)^{\mathfrak{S}_\infty}$ (called the **Ivanov-Kerov algebra**) has a linear basis labelled by partitions of arbitrary size:

$$\mathscr{B}_\infty = \operatorname{Vect}((A_\lambda)_{\lambda \in \mathfrak{Y}}), \quad \text{with } A_\mu = \sum_{|S|=|\mu|} \sum_{\substack{\sigma \in \mathfrak{S}(S) \\ t(\sigma)=\mu}} (\sigma, S).$$

Actually, we would rather use a renormalized version of this basis, defined as follows:

$$\Sigma_\mu = \sum \left[(a_{11}, \ldots, a_{1\mu_1})(a_{21}, \ldots, a_{2\mu_2}) \cdots (a_{r1}, \ldots, a_{r\mu_r}), \{a_{ij}\}_{\substack{1 \le i \le \ell(\mu) \\ 1 \le j \le \mu_i}} \right]$$

where the sum is taken over injective $\mathbb{N}^*$-valued functions $i, j \mapsto a_{ij}$ with $1 \le i \le \ell(\mu)$ and $1 \le j \le \mu_i$.

Proposition 9. *The family $(\Sigma_\mu)_{\mu \in \mathfrak{Y}}$ is a linear basis of $\mathscr{B}_\infty$, and for any n and any partition $\lambda \in \mathfrak{Y}_n$,*

$$\Sigma_\mu(\lambda) = \chi^\lambda(\pi_n(\Sigma_\mu)).$$

Proof. Here, an irreducible character χ^λ is extended by linearity from the group $\mathfrak{S}_n$ to the algebra $\mathbb{C}\mathfrak{S}_n$. It is not hard to see that Σ_μ is a multiple of A_μ for any partition μ; hence, $(\Sigma_\mu)_\mu$ is indeed a linear basis of $\mathscr{B}_\infty$. Now, for any n, $\pi_n(\Sigma_\mu)$ consists in $n^{\downarrow|\mu|}$ elements of $\mathfrak{S}_n$ of cycle type $\mu \sqcup 1^{n-|\mu|}$, corresponding to all the possible $|\mu|$-arrangements of the a_{ij}'s. Hence,

$$\chi^\lambda(\pi_n(\Sigma_\mu)) = n^{\downarrow|\mu|} \chi^\lambda(\mu \sqcup 1^{n-|\mu|}) = \Sigma_\mu(\lambda).$$

$\square$

Now, Proposition 9 allows to compute a product of two symbols Σ in $\mathscr{B}_\infty$ instead of $\mathscr{O}$. Indeed, for any partitions λ, μ, ν,

$$\Sigma_\mu(\lambda)\, \Sigma_\nu(\lambda) = \chi^\lambda(\pi_n(\Sigma_\mu))\, \chi^\lambda(\pi_n(\Sigma_\nu)) = \chi^\lambda(\pi_n(\Sigma_\mu)\, \pi_n(\Sigma_\nu))$$

$$= \chi^\lambda(\pi_n(\Sigma_\mu \times_{\mathscr{B}_\infty} \Sigma_\nu)) = (\Sigma_\mu \times_{\mathscr{B}_\infty} \Sigma_\nu)(\lambda)$$

where we used the fact that for any elements a, b of the center $Z(\mathbb{C}\mathfrak{S}_n)$ of the group algebra, and any normalized irreducible character χ^λ of $\mathfrak{S}_n$, $\chi^\lambda(ab) = \chi^\lambda(a)\,\chi^\lambda(b)$ — this is true for any finite group. With this new point of view, the combinatorics of the symbols Σ has been investigated by P. Śniady in [Śni06a, Śni06b], and later by V. Féray and the author in [FM10]. Thus, one can give a semi-explicit description of a product $\Sigma_\mu \Sigma_\nu$ by using **partial matchings**.

To this purpose, let us introduce some additional notations. We denote by $\mathfrak{A}_k$ the set of k-arrangements of integers in $\mathbb{N}^*$, and if $A = (a_1, \ldots, a_k) \in \mathfrak{A}_k$, we denote by $C(A)$ the corresponding k-cycle, viewed as a partial permutation with support $\{a_1, \ldots, a_k\}$. Then, Σ_μ can be rewritten as

$$\Sigma_\mu = \sum_{\substack{\forall i, \ A_i \in \mathfrak{A}_{\mu_i} \\ \forall i \neq j, \ A_i \cap A_j = \emptyset}} C(A_1)\, C(A_2) \cdots C(A_{\ell(\mu)}),$$

and given two partitions μ and ν,

$$\Sigma_\mu\, \Sigma_\nu = \sum_{\substack{\forall i, \ A_i \in \mathfrak{A}_{\mu_i} \\ \forall i \neq j, \ A_i \cap A_j = \emptyset}} \ \sum_{\substack{\forall k, \ B_k \in \mathfrak{A}_{\nu_k} \\ \forall k \neq l, \ B_k \cap B_l = \emptyset}} C(A_1) \cdots C(A_{\ell(\mu)})\, C(B_1) \cdots C(B_{\ell(\nu)}).$$

To any sequences $(a_{ij})_{i,j}$ and $(b_{kl})_{k,l}$ appearing in this expression, we associate a partial matching of the set I_A of the (i,j)'s with the set I_B of the (k,l)'s:

$$(i, j) \sim (k, l) \text{ if } a_{ij} = b_{kl}.$$

Then, given a partial matching M of I_A with I_B, if we sum the products

$$C(A_1) \cdots C(A_{\ell(\mu)})\, C(B_1) \cdots C(B_{\ell(\nu)})$$

corresponding to this partial matching, we obtain a term $\Sigma_{\rho(M,\mu,\nu)}$: indeed, if one knows which of the a_{ij}'s are equal to which of the b_{kl}'s, then one can rewrite the product as a product of *disjoint* cycles in a way that does not depend on the values of the a' and the b's (*cf.* [FM10, §3.4]). Thus:

Lemma 10. *For any partition μ and ν,*

$$\Sigma_\mu\, \Sigma_\nu = \sum_M \Sigma_{\rho(M,\mu,\nu)}$$

where the sum runs over all the partial matchings M of I_A with I_B.

We refer again to [FM10] for more details on this combinatorial lemma. The reader should also notice that the degree of a term $\Sigma_{\rho(M,\mu,\nu)}$ corresponding to a partial matching M is $|\mu| + |\nu| - |M|$, where $|M|$ is the size of the partial matching, that is to say the number of pairs $(a \in I_A, b \in I_B)$ in M.

Example. Let us compute the term of higher degree and the subdominant term of a product $\Sigma_\mu\, \Sigma_\nu$.

0. If $\deg \Sigma_{\rho(M,\mu,\nu)} = |\mu| + |\nu|$, then $|M| = 0$, and $M = \emptyset$. The cycles corresponding to the empty matching are disjoint ones; thus, the only term of degree $|\mu| + |\nu|$ in $\Sigma_\mu\, \Sigma_\nu$ is indeed $\Sigma_{\mu \sqcup \nu}$.

1. If $\deg \Sigma_{\rho(M,\mu,\nu)} = |\mu| + |\nu| - 1$, then $|M| = 1$ and M contains only one pair (a, b). Suppose that a is in a cycle $C(A_i)$ of length μ_i, and that b is in a cycle $C(B_j)$ of length ν_j. Then, $C(A_i)$ and $C(B_j)$ intersect in exactly one point, so their product is a $(\mu_i + \nu_j - 1)$-cycle; and all the other cycles $C(A_{i'})$ or $C(B_{j'})$ stay disjoint. Thus, such a matching yields a term

$$\Sigma_{(\mu \sqcup \nu \sqcup \{\mu_i + \nu_j - 1\}) \setminus \{\mu_i, \nu_j\}},$$

   and one has to sum over all possible matchings of size 1 to get the subdominant term of $\Sigma_\mu \Sigma_\nu$.

So, for any partitions $\mu = (\mu_1, \ldots, \mu_r)$ and $\nu = (\nu_1, \ldots, \nu_s)$,

$$\Sigma_\mu \, \Sigma_\nu = \Sigma_{\mu \sqcup \nu} + \sum_{i=1}^{r} \sum_{j=1}^{s} \mu_i \nu_j \, \Sigma_{(\mu \sqcup \nu \sqcup \{\mu_i + \nu_j - 1\}) \setminus \{\mu_i, \nu_j\}}$$

plus an observable of degree at most $|\mu| + |\nu| - 2$. In particular, for two symbols Σ_l and Σ_m corresponding to cycles,

$$\Sigma_l \, \Sigma_m = \Sigma_{l,m} + lm \, \Sigma_{l+m-1} + (\text{observable of degree at most } l + m - 2).$$

3.3 Śniady's theory of cumulants of observables

In order to prove results of asymptotic gaussian behaviour in the setting of the representation theory of $\mathfrak{S}_\infty$, we shall use in Section 5 Śniady's theory of **cumulants of observables** (*cf.* [Śni06a, Śni06b]), though in a setting that is slightly different from the setting of his papers, and with another gradation on the algebra $\mathscr{O}$. If $X_1, \ldots, X_r$ are real random variables defined on a common probability space, we recall that their **joint cumulant** is defined by:

$$k(X_1, X_2, \ldots, X_r) = \left. \frac{\partial^r}{\partial t_1 \partial t_2 \cdots \partial t_r} \right|_{t=0} \log \mathbb{E}\left[\exp(t_1 X_1 + t_2 X_2 + \cdots + t_r X_r)\right].$$

In particular, $k(X) = \mathbb{E}[X]$ is the expectation and $k(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y]$ is the covariance. With the definition given above, it is obvious that a gaussian vector $X = (X_1, \ldots, X_n)$ is characterized by the vanishing of all the joint cumulants of the coordinates of order $r \geq 3$. On the other hand, one can give an equivalent combinatorial definition of the joint cumulants. For any integer n, we denote by $\mathfrak{P}_n$ the set of **set partitions** of $[\![1, n]\!]$, and if $\pi = \pi_1 \sqcup \pi_2 \sqcup \cdots \sqcup \pi_s$ is such a set partition, we denote $\ell(\pi) = s$ and $\mu(\pi) = (-1)^{\ell(\pi)-1}\,(\ell(\pi) - 1)!$. Then,

$$k(X_1, X_2, \ldots, X_r) = \sum_{\pi \in \mathfrak{P}_r} \mu(\pi) \left(\prod_{i=1}^{\ell(\pi)} \mathbb{E}\left[\prod_{j \in \pi_i} X_j\right]\right).$$

Example. For instance, when $r = 3$, one has to sum over the 5 set partitions of $\mathfrak{P}_3$, and $k(X, Y, Z)$ is equal to

$$\mathbb{E}[XYZ] - \mathbb{E}[XY]\,\mathbb{E}[Z] - \mathbb{E}[XZ]\,\mathbb{E}[Y] - \mathbb{E}[YZ]\,\mathbb{E}[X] + 2\,\mathbb{E}[X]\,\mathbb{E}[Y]\,\mathbb{E}[Z].$$

Now, in the algebra of observables of diagrams $\mathscr{O}$, we can in fact define three kind of cumulants. First, if $\mathbb{P}$ is a probability law on partitions, then any $X \in \mathscr{O}$ can be considered as a random variable $X(\lambda)$, and one can define the joint cumulant of observables $X_1, \ldots, X_r$. That said, one can define on $\mathscr{O} \simeq \mathscr{B}_\infty$ a new product $\bullet$, called **disjoint product**, and given by the following rule:

$$\forall \mu, \nu, \ \ \Sigma_\mu \bullet \Sigma_\nu = \Sigma_{\mu \sqcup \nu}.$$

This new product is the restriction to $\mathscr{B}_\infty$ of the disjoint product of partial permutations:

$$(\sigma, S) \bullet (\tau, T) = \begin{cases} (\sigma\tau, S \sqcup T) & \text{if } S \cap T = \emptyset, \\ 0 & \text{otherwise.} \end{cases}$$

We will denote by $\mathscr{O}^\bullet$ the algebra of observables endowed with this new product. The **disjoint cumulant** of some observables $X_1, \ldots, X_r$ is then defined exactly in the same way as for the standard cumulant:

$$k^\bullet(X_1, X_2, \ldots, X_r) = \sum_{\pi \in \mathfrak{P}_r} \mu(\pi) \left(\prod_{i \leq \ell(\pi)} \mathbb{E}\left[\overset{\bullet}{\prod_{j \in \pi_i}} X_j \right] \right).$$

Finally, we define the **identity cumulant** of observables $X_1, \ldots, X_r$ as the new observable

$$k^{\mathrm{id}}(X_1, X_2, \ldots, X_r) = \sum_{\pi \in \mathfrak{P}_r} \mu(\pi) \left(\overset{\bullet}{\prod_{i \leq \ell(\pi)}} \left[\prod_{j \in \pi_i} X_j \right] \right).$$

The three kind of cumulants are related by the following identity, that can be understood as a computation of expectation *via* conditioning, see [Bri69] and [Śni06b, Lemma 4.8]:

$$k(X_1, \ldots, X_r) = \sum_{\pi \in \mathfrak{P}_r} k^\bullet \left(k^{\mathrm{id}}(X_{j \in \pi_1}), k^{\mathrm{id}}(X_{j \in \pi_2}), \ldots, k^{\mathrm{id}}(X_{j \in \pi_{\ell(\pi)}}) \right).$$

The expansion of an identity cumulant $k^{\mathrm{id}}(\Sigma_{l_1}, \ldots, \Sigma_{l_r})$ in partial permutations has been computed in [FM10, Lemma 19]. Given arrangements $A_1, \ldots, A_r$, we define a relation $\sim$ by $i \sim j \iff A_i \cap A_j \neq \emptyset$. The transitive cloture of $\sim$ is an equivalence relation, and we denote by $\pi(A_1, \ldots, A_r)$ the corresponding set partition in $\mathfrak{P}_r$.

Lemma 11. *For any $l_1, \ldots, l_r$,*

$$k^{\mathrm{id}}(\Sigma_{l_1}, \ldots, \Sigma_{l_r}) = \sum_{\substack{\forall i, \ A_i \in \mathfrak{A}_{l_i} \\ \pi(A_1, \ldots, A_r) = [\![1,r]\!]}} C(A_1)\, C(A_2) \cdots C(A_r).$$

Proof. The three kind of cumulants can be defined inductively by the relations:

$$X_1 X_2 \cdots X_r = \sum_{\pi \in \mathfrak{P}_r} k^{\mathrm{id}}(X_{j \in \pi_1}) \bullet k^{\mathrm{id}}(X_{j \in \pi_2}) \bullet \cdots \bullet k^{\mathrm{id}}(X_{j \in \pi_{\ell(\pi)}});$$

$$\mathbb{E}[X_1 X_2 \cdots X_r] = \sum_{\pi \in \mathfrak{P}_r} k(X_{j \in \pi_1})\, k(X_{j \in \pi_2}) \cdots k(X_{j \in \pi_{\ell(\pi)}});$$

$$\mathbb{E}[X_1 \bullet X_2 \bullet \cdots \bullet X_r] = \sum_{\pi \in \mathfrak{P}_r} k^{\bullet}(X_{j \in \pi_1})\, k^{\bullet}(X_{j \in \pi_2}) \cdots k^{\bullet}(X_{j \in \pi_{\ell(\pi)}}).$$

In particular, the first relation enables us to prove Proposition 11 by induction on r, see [FM10, Lemma 19]. $\qquad\square$

Now, the degree in $\mathscr{O} \simeq \mathscr{B}_\infty$ is the restriction of the degree $\deg(\sigma, S) = |S|$ of the algebra of partial permutations $\mathscr{A}_\infty$. However, in any product of cycles appearing in the expansion of $k^{\mathrm{id}}(\Sigma_{l_1}, \ldots, \Sigma_{l_r})$, the joint support of the cycles $C(A_1), \ldots, C(A_r)$ has for cardinality at most $l_1 + l_2 + \cdots + l_r - (r-1)$ — this is imposed by the condition $\pi(A_1, \ldots, A_r) = [\![1, r]\!]$. Hence, Proposition 11 implies that

$$\forall l_1, \ldots, l_r, \ \ \deg k^{\mathrm{id}}(\Sigma_{l_1}, \ldots, \Sigma_{l_r}) \leq l_1 + \cdots + l_r - (r-1).$$

This leads to the following estimate, that will be crucial in the proof of our central limit theorem for the irreducible characters of $\mathfrak{S}_\infty$.

Theorem 12. *For $\omega \in \Omega$, we denote by $k_{n,\omega}$ and $k_{n,\omega}^{\bullet}$ the cumulant and the disjoint cumulant of observables with respect to the expectation $\mathbb{E}_{n,\omega}$. For any observables $X_1, \ldots, X_r$,*

$$k_{n,\omega}^{\bullet}(X_1, \ldots, X_r) = O\left(n^{\deg X_1 + \cdots + \deg X_r - r + 1}\right);$$

$$k_{n,\omega}(X_1, \ldots, X_r) = O\left(n^{\deg X_1 + \cdots + \deg X_r - r + 1}\right).$$

Proof. Because of the multilinearity of cumulants and of some compatibility relation with respect to the products (see [Śni06b, Theorem 4.4]), it is sufficient to prove these estimates for an algebraic basis of $\mathscr{O}^{\bullet}$ or $\mathscr{O}$, so we will suppose in the following that $X_i = \Sigma_{l_i}$ with positive integers $l_1, \ldots, l_r \geq 1$. Now, notice that for any partition μ,

$$\mathbb{E}_{n,\omega}[\Sigma_\mu] = p_\mu(A - B + \gamma E)\, n^{\downarrow |\mu|} = p_\mu(A - B + \gamma E)\, \mathbb{E}_{n,\omega}[\Sigma_{1^{|\mu|}}].$$

Thus, one can rewrite a disjoint cumulant $k_{n,\omega}^{\bullet}(\Sigma_{l_1}, \ldots, \Sigma_{l_r})$ as follows:

$$k_{n,\omega}^{\bullet}(\Sigma_{l_1}, \ldots, \Sigma_{l_r}) = \sum_{\pi \in \mathfrak{P}_r} \mu(\pi) \left(\prod_{i \leq \ell(\pi)} \mathbb{E}_{n,\omega} \left[\prod_{j \in \pi_i}^{\bullet} \Sigma_{l_j} \right] \right)$$

$$= p_{l_1, \ldots, l_r}(A - B + \gamma E) \sum_{\pi \in \mathfrak{P}_r} \mu(\pi) \left(\prod_{i \leq \ell(\pi)} \mathbb{E}_{n,\omega} \left[\prod_{j \in \pi_i}^{\bullet} \Sigma_{1^{l_j}} \right] \right)$$

$$= p_{l_1, \ldots, l_r}(A - B + \gamma E) \, k_{n,\omega}^{\bullet}(\Sigma_{1^{l_1}}, \ldots, \Sigma_{1^{l_r}}).$$

So, in order to prove the estimate for disjoint cumulants, it is sufficient to treat the case of observables Σ_{1^l}. However, for these observables, the expectation $\mathbb{E}_{n,\omega}[\Sigma_{1^l}] = n^{\downarrow l}$ does not depend on ω, and the estimate is simply an algebraic result on the degree of some polynomial combinations of falling factorials of n. It has been proved in [Śni06b, Lemma 4.8] that:

$$k_n^{\bullet}(\Sigma_{1^{l_1}}, \ldots, \Sigma_{1^{l_r}}) = O\left(n^{l_1 + \cdots + l_r - r + 1} \right).$$

Thus, the proof is complete for the disjoint cumulants. As for the standard cumulants, one uses the aforementioned formula:

$$k_{n,\omega}(X_1, \ldots, X_r) = \sum_{\pi \in \mathfrak{P}_r} k_{n,\omega}^{\bullet}\left(k^{\mathrm{id}}(X_{j \in \pi_1}), k^{\mathrm{id}}(X_{j \in \pi_2}), \ldots, k^{\mathrm{id}}(X_{j \in \pi_{\ell(\pi)}}) \right).$$

Since the estimate for disjoint cumulants has been proved, it suffices now to show that for any set partition $\pi \in \mathfrak{P}_r$,

$$\left(\sum_{i=1}^{\ell(\pi)} \deg k^{\mathrm{id}}(X_j \in \pi_i) \right) - \ell(\pi) + 1 \leq \deg X_1 + \cdots + \deg X_r - r + 1.$$

This is trivial because of Proposition 11:

$$\sum_{i=1}^{\ell(\pi)} \deg k^{\mathrm{id}}(X_j \in \pi_i) \leq \sum_{i=1}^{\ell(\pi)} \left(\sum_{j \in \pi_i} \deg X_i - |\pi_i| + 1 \right)$$

$$\leq \deg X_1 + \cdots + \deg X_r - r + \ell(\pi).$$

Notice that the estimates of our Theorem can be made independent of ω: indeed, for any integers $l_1, \ldots, l_r$, $|p_{l_1, \ldots, l_r}(A - B + \gamma E)| \leq 1$, and these quantities are the only ones that depend explicitly on ω in our proof. $\qquad\square$

A particular case of Theorem 12 is when $r = 1$; thus, $\mathbb{E}_{n,\omega}[X] = O(n^{\deg X})$ for any observable X. This simple fact and the multiplicativity of the irreducible characters of $\mathfrak{S}_\infty$ will now be used in Section 4, in which we shall recall the results of Kerov and Vershik (see [KV81a]) and their proofs.

4 Kerov-Vershik law of large numbers

In this section, we recall the results of Kerov and Vershik; although nothing is new here, we shall introduce the important concept of a (random) probability measure associated to a (random) diagram, that will later prove useful for the analysis of the fluctuations of the rows and columns, see §5.2.

4.1 Law of large numbers for the irreducible characters

Given an irreducible character χ^ω, we replace it on $\mathfrak{S}_n$ by the random function χ^λ, where λ is taken according to the probability measure $\mathbb{P}_{n,\omega}$. The law of large numbers of Kerov and Vershik ensures that χ^λ converges to χ^ω:

Theorem 13. *For any permutation $\sigma \in \mathfrak{S}_\infty$,*

$$\chi^\lambda(\sigma) \longrightarrow_{\mathbb{P}_{n,\omega}} \chi^\omega(\sigma),$$

where the long right arrow means that one has convergence in probability.

Proof. For any partition μ, $\mathbb{E}_{n,\omega}\left[\frac{\Sigma_\mu}{n^{|\mu|}}\right] \simeq p_\mu(A - B + \gamma E)$, and

$$\mathbb{E}_{n,\omega}\left[\left(\frac{\Sigma_\mu}{n^{|\mu|}}\right)^2\right] = \mathbb{E}_{n,\omega}\left[\frac{\Sigma_{\mu\sqcup\mu}}{n^{2|\mu|}}\right] + \frac{1}{n^{2|\mu|}}\,\mathbb{E}_{n,\omega}[\text{obs. of degree at most } 2|\mu| - 1]$$

$$= (p_\mu(A - B + \gamma E))^2 + O(n^{-1}) \simeq (p_\mu(A - B + \gamma E))^2.$$

Here, we have used the estimate $\mathbb{E}_{n,\omega}[X] = O(n^{\deg X})$ mentioned at the end of the previous paragraph, the factorization property of the symbols Σ_μ in higher degree, and the computation of the expectations $\mathbb{E}_{n,\omega}[\Sigma_\mu]$ performed at the end of §3.1. Because of Bienaymé-Chebyshev inequality, these two estimates ensure that $\frac{\Sigma_\mu}{n^{|\mu|}} \simeq \chi^\lambda(\mu \sqcup 1^{n-|\mu|})$ converges in probability towards $p_\mu(A - B + \gamma E) = \chi^\omega(\mu)$. $\qquad\square$

Again, Theorem 13 can be given an uniform flavour: namely, for any $\sigma \in \mathfrak{S}_\infty$, and any $\varepsilon > 0$,

$$\sup_{\omega \in \Omega} \mathbb{P}_{n,\omega}[|\chi^\lambda(\sigma) - \chi^\omega(\sigma)| \geq \varepsilon] \to_{n\to\infty} 0.$$

On the other hand, we shall see in Section 5 that in fact, a much stronger result holds: for any permutation σ, the renormalized difference $\sqrt{n}\,(\chi^\lambda(\sigma) - \chi^\omega(\sigma))$ converges in law towards a centered gaussian variable. Now, let us remark that for any k,

$$\mathbb{E}_{n,\omega}\left[\left(\frac{p_k - \Sigma_k}{n^k}\right)^2\right] = \frac{\mathbb{E}_{n,\omega}[\text{obs. of degree at most } 2k - 2]}{n^{2k}} = O(n^{-2}) \to 0.$$

So, Theorem 13 implies that for any $k \geq 1$, and any $\omega \in \Omega$,

$$\frac{p_k(\lambda)}{n^k} \longrightarrow_{\mathbb{P}_{n,\omega}} p_k(A - B + \gamma E).$$

We are now going to interpret this result in geometric terms, thereby proving a law of large numbers for the first rows $\lambda_1, \lambda_2, \ldots$ and the first columns $\lambda'_1, \lambda'_2, \ldots$ of a random partition λ chosen according to a law $\mathbb{P}_{n,\omega}$.

4.2 Random measures associated to random partitions

We denote by $\mathscr{P}([-1,1])$ the set of borelian probability measures on the interval $[-1, 1]$; this is a metrizable compact set for Skorohod's topology of weak convergence, see [Bil69, Chapter 1]. A distance compatible with this topology is provided by any dense sequence of functions in the space of continuous (real-valued) functions $\mathscr{C}([-1,1])$ — by Stone-Weierstrass theorem, such a sequence $(f_n)_{n \in \mathbb{N}}$ exists, and we can for instance take $\{f_n\}_{n \in \mathbb{N}} = \mathbb{Q}[x]$. Thus,

$$d(m_1, m_2) = \sum_{n=0}^{\infty} \frac{1}{2^n} \min(1, |m_1(f_n) - m_2(f_n)|)$$

is a distance on $\mathscr{P}([-1,1])$ that defines the topology of weak convergence[6]. Now, to any partition λ of size n, we associate a probability measure $X_\lambda \in \mathscr{P}([-1,1])$:

$$X_\lambda = \sum_{i=1}^{d} \frac{a_i}{n} \, \delta_{\left(\frac{a_i}{n}\right)} + \sum_{i=1}^{d} \frac{b_i}{n} \, \delta_{\left(-\frac{b_i}{n}\right)},$$

where the a's and the b's are the Frobenius coordinates of λ. Notice that the moments of X_λ are up to a shift of index the renormalized Frobenius moments of λ:

$$\forall k \geq 0, \ \ X_\lambda(x^k) = \frac{p_{k+1}(\lambda)}{n^{k+1}}.$$

If λ is random, we will then consider X_λ as a **random measure**, that is to say a random point of the space $\mathscr{P}([-1,1])$.

If $\omega = ((\alpha_1, \alpha_2, \ldots), (\beta_1, \beta_2, \ldots))$ is a point of the Thoma simplex Ω, we can also associate to it a probability measure $X_\omega \in \mathscr{P}([-1,1])$:

$$X_\omega = \sum_{i=1}^{\infty} \alpha_i \, \delta_{(\alpha_i)} + \sum_{i=1}^{\infty} \beta_i \, \delta_{(-\beta_i)} + \gamma \, \delta_0.$$

The moments of X_ω are up to a shift of index the specializations of the power sums on the formal alphabet $A - B + \gamma E$:

$$\forall k \geq 0, \ \ X_\omega(x^k) = p_{k+1}(A - B + \gamma E).$$

---

[6] Here, for any $f \in \mathscr{C}([-1,1])$ and any $m \in \mathscr{P}([-1,1])$, $m(f) = \int_{-1}^{1} f(x) \, m(dx)$.

We have seen in the previous paragraph that under the coherent system of measures $\mathbb{P}_{n,\omega}$, $\frac{p_k(\lambda)}{n^k}$ converges in probability towards $p_k(A - B + \gamma E)$ for any k. Consequently,

$$\forall k \geq 0, \quad X_\lambda(x^k) \longrightarrow_{\mathbb{P}_{n,\omega}} X_\omega(x^k).$$

By linearity, we have the same result for any evaluation of the measures on a polynomial in $\mathbb{Q}[x]$, and since $\mathbb{Q}[x]$ is dense in $\mathscr{C}([-1,1])$, we conclude that:

Proposition 14. *For the topology of weak convergence,*

$$X_\lambda \longrightarrow_{\mathbb{P}_{n,\omega}} X_\omega.$$

More concretely, for *any* function $f \in \mathscr{C}([-1,1])$ and any $\varepsilon > 0$, $\mathbb{P}_{n,\omega}[|X_\lambda(f) - X_\omega(f)| \geq \varepsilon] \to 0$. In fact, these convergences are again uniform on the Thoma simplex:

$$\sup_{\omega \in \Omega} \mathbb{P}_{n,\omega}[|X_\lambda(f) - X_\omega(f)| \geq \varepsilon] \to_{n \to \infty} 0.$$

Finally, let us translate the Kerov-Vershik law of large numbers in terms of the asymptotic geometry of the random partitions. From the result on random measures, it is easy to guess that the following should be true:

Theorem 15. *Take λ randomly according to a measure $\mathbb{P}_{n,\omega}$, with $\omega = (\alpha, \beta)$.*

$$\forall i \geq 1, \quad \frac{\lambda_i}{n} \simeq \frac{a_i}{n} \longrightarrow_{\mathbb{P}_{n,\omega}} \alpha_i \qquad ; \qquad \forall j \geq 1, \quad \frac{\lambda'_j}{n} \simeq \frac{b_j}{n} \longrightarrow_{\mathbb{P}_{n,\omega}} \beta_j.$$

Proof. One of the main difficulty when recovering the shapes of the partitions is that we have to take account of the multiplicities of the coordinates α_i, β_j of the point $\omega \in \Omega$. Let us prove rigorously that if $\alpha_1 = \alpha_2 = \cdots = \alpha_m > \alpha_{m+1}$ are the biggest coordinates of α, then the renormalized m first parts of λ converge in probability to the value α_1. We fix $\varepsilon > 0$, and we split the proof in two parts.

1. First, let us prove that $\frac{a_1}{n}$ is smaller than $\alpha_1 + 2\varepsilon$ with probability going to 1 as n goes to infinity (of course, we suppose $\alpha_1 < 1$, this case being obvious). We consider a test function f_ε^+ that takes its values in $[0,1]$, is continuous, and is equal to 1 after $\alpha_1 + 2\varepsilon$ and to 0 before $\alpha_1 + \varepsilon$ (see Figure 1). On the one hand, $X_\omega(f_\varepsilon^+) = 0$, and on the other hand,

$$X_\lambda(f_\varepsilon^+) \geq (\alpha_1 + 2\varepsilon) \operatorname{card} \left\{ i \mid \frac{a_i}{n} \geq \alpha_1 + 2\varepsilon \right\}.$$

Since $X_\lambda(f_\varepsilon)$ converges in probability to $X_\omega(f_\varepsilon^+) = 0$, this implies that the cardinality is equal to 0 for n big enough and with big probability.
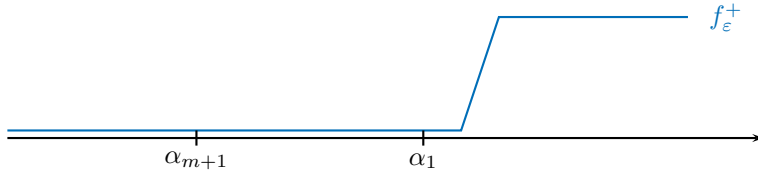
Fig. 1 Test function f_ε^+.
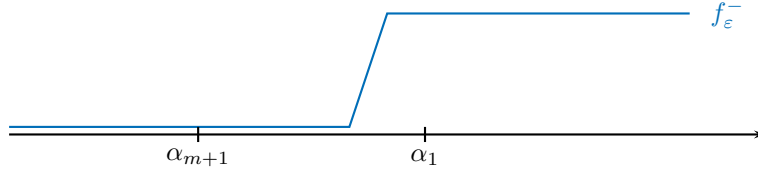

Fig. 2 Test function f_ε^-.

2. Similarly, let us prove that $\frac{a_m}{n}$ is bigger than $\alpha_1 - 2\varepsilon$ with probability going to 1 as n goes to infinity (again, we suppose that $\alpha_1 > 0$, this case being obvious). We consider a test function f_ε^- that takes its values in $[0, 1]$, is continuous, and is equal to 1 after $\alpha_1 - \varepsilon$ and to 0 before $\alpha_1 - 2\varepsilon$ (see Figure 2). For ε small enough, we may also suppose that $\alpha_1 - 2\varepsilon > \alpha_{m+1}$. Then, $X_\omega(f_\varepsilon^-) = m\alpha_1$, whereas

$$X_\lambda(f_\varepsilon^-) \le (\alpha_1 + 2\varepsilon) \, \text{card} \left\{ i \mid \frac{a_i}{n} \ge \alpha_1 - 2\varepsilon \right\}$$

with big probability — indeed, we already know that the a_i's are smaller than $\alpha_1 + 2\varepsilon$ with probability almost 1. For ε small enough and as n goes to infinity, this leads to the inequality $\text{card} \{ i \mid \frac{a_i}{n} \ge \alpha_1 - 2\varepsilon \} \ge m$, which is equivalent to ask that $\frac{a_m}{n} \ge \alpha_1 - 2\varepsilon$.

Thus, for any $\varepsilon > 0$, the probability of the event

$$\alpha_1 + 2\varepsilon \ge \frac{a_1}{n} \ge \cdots \ge \frac{a_m}{n} \ge \alpha_1 - 2\varepsilon$$

goes to 1 as n goes to infinity, which is exactly Theorem 15 for the m first rows of λ. Now, an induction enables us to treat the other rows, and the columns are treated exactly the same way. □

Example. When $\omega = \mathbf{0}$, the law of large numbers ensures that all the rows of λ are some $o(n)$, and similarly for its columns. Actually, it is a well-known fact that the rows and the columns have order of magnitude $\sqrt{n}$, and that there is a limit shape for the random partitions under Plancherel measures after scaling the Young diagrams in both directions by the factor $\sqrt{n}$ (*cf.* [LS77, KV77]):
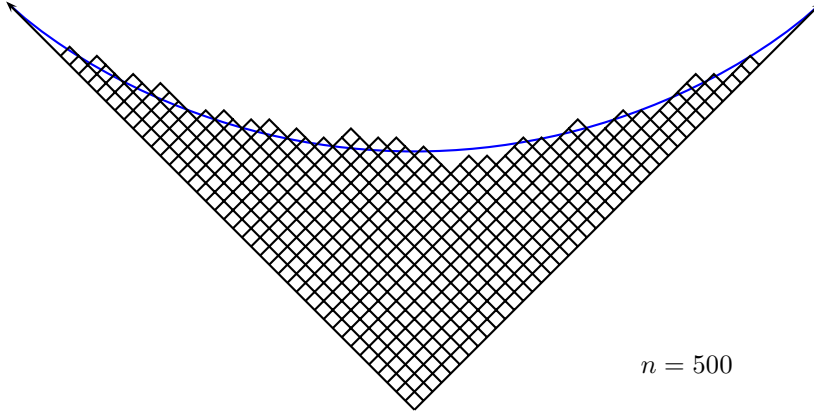
$$n = 500$$

Fig. 3 Random partition under the Plancherel measure $\mathbb{P}_{n,\mathbf{0}} = \mathbb{P}_n$.

Hence, for this particular point of Ω, the law of large numbers is not precise enough; however, an accurate law of large numbers and a central limit theorem can be proved in this case (and for other classes of measures on partitions) by a slight modification of the tools that we use here (see [IO02, Mél11]). Another important example is when $\omega = (\alpha, \beta)$ with α and β geometric progressions of same ratio q. However, these examples are better understood in the setting of the Hecke algebras, see Section 7.

To conclude this section, let us interpret the geometric law of large numbers in terms of properties of ω-shuffles. We fix an integer r. Given a shuffle σ of $n_1 + n_2 + \cdots + m_1 + m_2 + \cdots + l$ cards as in §2.3, it is obvious that the maximal sum

$$\ell(iw_1) + \ell(iw_2) + \cdots + \ell(iw_r)$$

of the lengths of r disjoint increasing subwords in σ is at least $n_1 + n_2 + \cdots + n_r$. Indeed, it suffices to consider the words

$$iw_j = [\![n_1 + \cdots + n_{j-1} + 1, n_1 + \cdots + n_j]\!],$$

because they stay in this order in σ.

Example. Take $r = 1$, $n = 9 = 5 + 4$. Any shuffle of 12345 and 6789 contains an increasing subword of length at least 5, namely, 12345. However, for many shuffles, one can do better: for instance, 162378495 is a shuffle of 12345 with 6789, and it contains an increasing subword of length 6, namely, 123789.

Thus, the r-th Greene invariant of σ, which we shall denote by $GI_r(\sigma)$, is bigger than $n_1 + \cdots + n_r$. Now, if the sizes of the blocks are taken according to a multinomial law of parameter (α, β, γ), then the law of large numbers

ensures that $\frac{n_i}{n}$ converges in probability to α_i. Thus, with probability going to 1, if σ is a permutation obtained by an ω-shuffle, then

$$\frac{GI_r(\sigma)}{n} \geq \alpha_1 + \cdots + \alpha_r - \varepsilon$$

for any $\varepsilon > 0$. However, one could imagine that "many" ω-shuffles yield permutations with greater Greene invariants (this impression may be strengthened by the randomized block of size $l \simeq n\gamma$). The Kerov-Vershik law of large numbers and Propositions 6 and 7 ensure that it is not the case:

Proposition 16. *Let σ be a permutation of size n chosen randomly according to $\mathbb{Q}_{n,\omega}$. As n goes to infinity,*

$$\frac{GI_r(\sigma)}{n} \longrightarrow_{\mathbb{Q}_{n,\omega}} \alpha_1 + \cdots + \alpha_r.$$

Of course, the same holds if one replaces the increasing subwords by the decreasing subwords and the α's by the β's.

Proof. Because of Proposition 6, the distribution of $GI_r(\sigma)$ under $\mathbb{Q}_{n,\omega}$ is the same as the distribution of $\lambda_1 + \cdots + \lambda_r$ under $\mathbb{P}_{n,\omega}$. Then, it suffices to apply Theorem 15. $\square$

Hence, ω-shuffles do not create bigger increasing subwords than those given by the sizes of the blocks. In Section 5, we are going to prove a central limit theorem that will improve this result. Hence, assuming that the α_i's are all distinct, up to order $\sqrt{n}$ (and not n), the asymptotic distribution of $GI_r(\sigma)$ is the same as the asymptotic distribution of the size $n_1 + \cdots + n_r$ of the r first blocks of the ω-shuffle, that is to say

$$n\left(\alpha_1 + \cdots + \alpha_r\right) + \sqrt{n}\,\sqrt{(\alpha_1 + \cdots + \alpha_r)(1 - \alpha_1 - \cdots - \alpha_r)}\,\mathcal{N}(0,1) + o(\sqrt{n})$$

with $\mathcal{N}(0,1)$ a standard gaussian variable.

5 Central limit theorems for characters and for partitions

In the previous paragraph, it was shown that all the coherent systems of measures $(\mathbb{P}_{n,\omega})_{n\in\mathbb{N}}$ have a property of concentration, that can be seen at the level of characters or directly for the geometry of the random partitions. In this section, we shall prove that the concentration is always gaussian, which is quite a striking fact. However, the geometric version of our central limit theorem will be a little less precise than the algebraic one: indeed, one cannot study with our method of moments the fluctuations of rows or columns corresponding to equal coordinates of the point ω of the Thoma simplex. This will be done in §6 by entirely different methods, that rely essentially on the interpretation of the measures $\mathbb{P}_{n,\omega}$ as push-forwards of the measures $\mathbb{Q}_{n,\omega}$.

5.1 Central limit theorem for the irreducible characters

In this paragraph, we fix permutations $\sigma^{(1)}, \ldots, \sigma^{(r)} \in \mathfrak{S}_\infty$ of cycle types $\mu^{(1)}, \ldots, \mu^{(r)}$; the partitions $\mu^{(i)}$ are all defined up to parts of size 1, but this will not change our results. For convenience, we shall denote $p_\mu(A - B + \gamma E) = p_\mu(\omega)$.

Theorem 17. *We consider the random vector* $(\Delta_{n,\omega}(\sigma^{(1)}), \ldots, \Delta_{n,\omega}(\sigma^{(r)}))$, *with*

$$\Delta_{n,\omega}(\sigma) = \sqrt{n} \left(\chi^\lambda(\sigma) - \chi^\omega(\sigma) \right)$$

and λ chosen randomly according to $\mathbb{P}_{n,\omega}$. As n goes to infinity, this vector converges in joint law to a centered gaussian vector $(\Delta_{\infty,\omega}(\sigma^{(1)}), \ldots, \Delta_{\infty,\omega}(\sigma^{(r)}))$, *and the covariance matrix of this vector is:*

$$
\begin{aligned}
& k(\Delta_{\infty,\omega}(\sigma^{(i)}), \Delta_{\infty,\omega}(\sigma^{(j)})) \\
& = \sum_{a \in \mu^{(i)}, \, b \in \mu^{(j)}} ab \, p_{\mu_i \sqcup \mu_j \setminus \{a,b\}}(\omega) \, (p_{a+b-1}(\omega) - p_a(\omega) \, p_b(\omega)).
\end{aligned}
$$

In particular, $k(\Delta_{\infty,\omega}(\sigma), \Delta_{\infty,\omega}(\tau)) = lm \, (p_{l+m-1}(\omega) - p_l(\omega) \, p_m(\omega))$ if σ and τ are cycles of lengths l and m.

Proof. For any partition μ,

$$\sqrt{n} \left(\chi^\lambda(\mu \sqcup 1^{n-|\mu|}) - p_\mu(\omega) \right) = \sqrt{n} \left(\frac{\Sigma_\mu(\lambda) - \mathbb{E}_{n,\omega}[\Sigma_\mu]}{n^{|\mu|}} \right) + O\left(\frac{1}{\sqrt{n}} \right),$$

so it is sufficient to show that the random vector

$$V = \sqrt{n} \left(\frac{\Sigma_{\mu^{(1)}}}{n^{|\mu^{(1)}|}}, \ldots, \frac{\Sigma_{\mu^{(r)}}}{n^{|\mu^{(r)}|}} \right)$$

converges to a gaussian vector with the announced covariance matrix. However, for $s \geq 3$, by using Theorem 12, one sees that

$$k_{n,\omega} \left(\Sigma_{\mu^{(i_1)}}, \ldots, \Sigma_{\mu^{(i_s)}} \right) = O\left(n^{|\mu^{(i_1)}| + \cdots + |\mu^{(i_s)}| - s + 1} \right)$$

$$k_{n,\omega} \left(\sqrt{n} \frac{\Sigma_{\mu^{(i_1)}}}{n^{|\mu^{(i_1)}|}}, \ldots, \sqrt{n} \frac{\Sigma_{\mu^{(i_s)}}}{n^{|\mu^{(i_s)}|}} \right) = O\left(n^{1 - \frac{s}{2}} \right) \to 0.$$

Thus, V converges indeed towards a gaussian vector. To compute the covariances, we use the following trick:

$$
\begin{aligned}
k_{n,\omega}(\Sigma_\mu, \Sigma_\nu) &= \mathbb{E}_{n,\omega}[\Sigma_\mu \, \Sigma_\nu] - \mathbb{E}_{n,\omega}[\Sigma_\mu] \, \mathbb{E}_{n,\omega}[\Sigma_\nu] \\
&= \mathbb{E}_{n,\omega}[\Sigma_\mu \, \Sigma_\nu - \Sigma_{\mu \sqcup \nu}] + \left(\mathbb{E}_{n,\omega}[\Sigma_{\mu \sqcup \nu}] - \mathbb{E}_{n,\omega}[\Sigma_\mu] \, \mathbb{E}_{n,\omega}[\Sigma_\nu] \right).
\end{aligned}
$$

The term of higher degree in the first part of the right-hand side is exactly the expectation of the subdominant term of a product $\Sigma_\mu \Sigma_\nu$. We have calculated this subdominant term after Proposition 10; hence,

$$\mathbb{E}_{n,\omega}[\Sigma_\mu \Sigma_\nu - \Sigma_{\mu \sqcup \nu}] = \sum_{a \in \mu, b \in \nu} ab \, \mathbb{E}_{n,\omega}[\Sigma_{(\mu \sqcup \nu \sqcup \{a+b-1\}) \setminus \{a,b\}}]$$

$$\simeq \left(\sum_{a \in \mu, \, b \in \nu} ab \, p_{\mu \sqcup \nu \setminus \{a,b\}}(\omega) \, p_{a+b-1}(\omega) \right) n^{|\mu|+|\nu|-1}.$$

On the other hand, $\mathbb{E}_{n,\omega}[\Sigma_{\mu \sqcup \nu}] - \mathbb{E}_{n,\omega}[\Sigma_\mu] \, \mathbb{E}_{n,\omega}[\Sigma_\nu]$ is easily computed to be equal to

$$p_{\mu \sqcup \nu}(\omega) \left[n^{\downarrow |\mu|+|\nu|} - n^{\downarrow |\mu|} \, n^{\downarrow |\nu|} \right] \simeq - |\mu| \, |\nu| \, p_{\mu \sqcup \nu}(\omega) \, n^{|\mu|+|\nu|-1}$$

Dividing everything by $n^{|\mu|+|\nu|-1}$ and writing

$$|\mu| \, |\nu| \, p_{\mu \sqcup \nu}(\omega) = \sum_{a \in \mu, \, b \in \nu} ab \, p_{\mu \sqcup \nu \setminus \{a,b\}}(\omega) \, p_a(\omega) \, p_b(\omega),$$

we obtain exactly the announced formula. $\square$

So, for any integer $k \geq 2$, if σ is a k-cycle, then

$$\sqrt{n} \left(\frac{\chi^\lambda(\sigma) - \chi^\omega(\sigma)}{k} \right)$$

converges under the laws $\mathbb{P}_{n,\omega}$ to a centered gaussian variable of variance $p_{2k-1}(\omega) - p_k(\omega)^2$. Again, this result is not precise enough when $\omega = \mathbf{0}$ — fortunately, this is the only problematic point. Actually, it has been shown in [IO02, Theorem 6.1] that under Plancherel measures, if σ is a k-cycle, then

$$n^{\frac{k}{2}} \left(\frac{\chi^\lambda(\sigma)}{\sqrt{k}} \right)$$

converges to a centered gaussian variable of variance 1. We shall say a little bit more on these kind of results in the conclusion of our paper.

5.2 Gaussian behaviour of the rows and of the columns

Now, let us look at the consequences of Theorem 17 for the geometry of the random partitions. First, let us give a central limit theorem for the random measures X_λ under coherent systems of probability measures $(\mathbb{P}_{n,\omega})_{n \in \mathbb{N}}$. In the following, $\mathscr{C}^1([-1,1])$ is endowed with the topology of uniform convergence of the functions and of their derivatives, and measures are now considered as elements of the dual space of $\mathscr{C}^1([-1,1])$ — this dual space contains $\mathscr{P}([-1,1]) \subset \mathscr{M}([-1,1]) = \mathscr{C}^0([-1,1])^*$.

Theorem 18. *We consider the random measure* $\nabla_{n,\omega} = \sqrt{n}\,(X_\lambda - X_\omega)$, *with* λ *chosen randomly according to* $\mathbb{P}_{n,\omega}$. *As* n *goes to infinity,* $\nabla_{n,\omega}$ *converges to a centered gaussian element* $\nabla_{\infty,\omega}$ *in* $\mathscr{C}^1([-1,1])^*$, *and for any functions* $f,g \in \mathscr{C}^1([-1,1])$,

$$k(\nabla_{\infty,\omega}(f), \nabla_{\infty,\omega}(g)) = X_\omega((xf(x))'(xg(x))') - X_\omega((xf(x))')\,X_\omega((xg(x))').$$

Proof. We have shown in our "algebraic" central limit theorem that the random vector

$$\left(\sqrt{n}\,\left(\frac{\Sigma_k}{n^k} - p_k(\omega) \right) \right)_{k \geq 1}$$

converges to a centered gaussian vector with covariance matrix

$$lm\,(p_{l+m-1}(\omega) - p_l(\omega)\,p_m(\omega)).$$

By using the same arguments as before, one can replace Σ_k by p_k, and therefore, the same results holds for

$$\left(\sqrt{n}\,\left(\frac{p_k(\lambda)}{n^k} - p_k(\omega) \right) \right)_{k \geq 1} = \left(\sqrt{n}\,\left(X_\lambda(x^k) - X_\omega(x^k) \right) \right)_{k \geq 0}$$
$$= (\nabla_{n,\omega}(x^k))_{k \geq 0},$$

although with a shift of index:

$$k(\nabla_{\infty,\omega}(x^l), \nabla_{\infty,\omega}(x^m)) = (l+1)(m+1)\,(p_{l+m+1}(\omega) - p_{l+1}(\omega)\,p_{m+1}(\omega)).$$

Now, this computation of covariances can be rewritten as:

$$k(\nabla_{\infty,\omega}(x^l), \nabla_{\infty,\omega}(x^m)) = (l+1)(m+1)\,(X_\omega(x^l\,x^m) - X_\omega(x^l)\,X_\omega(x^m))$$
$$= X_\omega((x^{l+1})'(x^{m+1})') - X_\omega((x^{l+1})')\,X_\omega((x^{m+1})').$$

By multilinearity, for any polynomials f and g in $\mathbb{R}[x]$, $\nabla_{n,\omega}(f)$ and $\nabla_{n,\omega}(g)$ converge towards centered gaussian variables, and the covariance of the limiting distibutions writes as

$$k(\nabla_{\infty,\omega}(f), \nabla_{\infty,\omega}(g)) = X_\omega((xf(x))'(xg(x))') - X_\omega((xf(x))')\,X_\omega((xg(x))').$$

Finally, one uses the density of $\mathbb{R}[x]$ in $\mathscr{C}^1([-1,1])$ to get the same result for any functions $f,g \in \mathscr{C}^1([-1,1])$. However, notice that one has first to apply a random version of the Banach-Steinhaus theorem for the (random) linear forms $\nabla_{n,\omega} : \mathscr{C}^1([-1,1]) \to \mathbb{R}$ (see [Tha87] or [VV95]). Indeed, because of the factor $\sqrt{n}$, it is not clear *a priori* that one can get a uniform bound on expressions like

$$\sqrt{n}\,(X_\lambda(f - f_k) - X_\omega(f - f_k))$$

with $(f_k)_{k \in \mathbb{N}}$ sequence of polynomials approximating f in $\mathscr{C}^1([-1,1])$ — this point was left unexplained in [FM10]. $\qquad\square$

Unfortunately, Theorem 18 is sufficient to describe the fluctuations of the row λ_i or the column λ'_j only when $\alpha_{i-1} > \alpha_i > \alpha_{i+1}$ or when $\beta_{j-1} > \beta_j > \beta_{j+1}$. Indeed, our method of moments will not allow to distinguish at order $\sqrt{n}$ the parts corresponding to equal coordinates. We fix distinct coordinates $\alpha_{i_1} \neq \alpha_{i_2} \neq \cdots \neq \alpha_{i_r} \neq 0$ and $\beta_{j_1} \neq \beta_{j_2} \neq \cdots \neq \beta_{j_s} \neq 0$ of the point $\omega \in \Omega$, and we denote by m_k (respectively, w_k) the multiplicity of α_{i_k} (resp., of β_{j_k}), so

$$\alpha_{i_k-1} > \alpha_{i_k} = \alpha_{i_k+1} = \cdots = \alpha_{i_k+m_k-1} > \alpha_{i_k+m_k}$$

and similarly for β_{j_k}. We then denote

$$\Lambda_{i_k,m_k,n} = \sqrt{n}\left(\frac{\lambda_{i_k}}{n} + \frac{\lambda_{i_k+1}}{n} + \cdots + \frac{\lambda_{i_k+m_k-1}}{n} - m_k\,\alpha_{i_k}\right);$$

$$\Lambda'_{j_k,w_k,n} = \sqrt{n}\left(\frac{\lambda'_{j_k}}{n} + \frac{\lambda'_{j_k+1}}{n} + \cdots + \frac{\lambda'_{j_k+w_k-1}}{n} - w_k\,\beta_{j_k}\right).$$

Theorem 19. *As n goes to infinity, the random vector*

$$(\Lambda_{i_1,m_1,n}, \ldots, \Lambda_{i_r,m_r,n}, \Lambda'_{j_1,w_1,n}, \ldots, \Lambda'_{j_s,w_s,n})$$

converges to a centered gaussian vector $(\Lambda_{i_1,m_1,\infty}, \ldots, \Lambda'_{j_s,l_s,\infty})$, with the following covariances:

$$k(\Lambda_{i_k,m_k,\infty}, \Lambda_{i_l,m_l,\infty}) = \delta_{kl}\,m_k\,\alpha_{i_k} - m_k m_l\,\alpha_{i_k}\alpha_{i_l};$$
$$k(\Lambda'_{j_k,w_k,\infty}, \Lambda'_{j_l,w_l,\infty}) = \delta_{kl}\,w_k\,\beta_{j_k} - w_k w_l\,\beta_{j_k}\beta_{j_l};$$
$$k(\Lambda_{i_k,m_k,\infty}, \Lambda'_{j_l,w_l,\infty}) = -m_k w_l\,\alpha_{i_k}\beta_{j_l}.$$

In particular, if α_i and α_j are two "isolated" coordinates of the sequence α, then the random variables $X_{i,n} = \sqrt{n}\left(\frac{\lambda_i}{n} - \alpha_i\right)$ and $X_{j,n} = \sqrt{n}\left(\frac{\lambda_j}{n} - \alpha_j\right)$ converge jointly in law to centered gaussian variables $X_{i,\infty}$ and $X_{j,\infty}$, with

$$k(X_{i,\infty}, X_{j,\infty}) = \delta_{ij}\,\alpha_i - \alpha_i\alpha_j.$$

Proof. We consider test functions f_k and g_l such that:

1. All these functions are in $\mathscr{C}^1([-1,1])$, non-negative and with values in $[0,1]$.

2. For any k, there is a neighbourhood $W_k = [\alpha_{i_k}-\varepsilon_k, \alpha_{i_k}+\varepsilon_k]$ of α_{i_k} such that W_k does not contain any element of the set $\{\alpha_{i\geq 1}, -\beta_{j\geq 1}, 0\} \setminus \{\alpha_{i_k}\}$, and the support of f_k is contained in W_k. Moreover, there is a neighboorhood $V_k \subset W_k$ of α_{i_k} such that $f_{k|V_k} = 1$.

3. Similarly, for any l, there is a neighbourhood $W'_l = [-\beta_{j_l} - \varepsilon'_l, -\beta_{j_l} + \varepsilon'_l]$ of $-\beta_{j_l}$ such that W'_l does not contain any element of $\{\alpha_{i\geq 1}, -\beta_{j\geq 1}, 0\} \setminus \{\beta_{j_l}\}$, and the support of g_l is contained in W'_l. Moreover, there is a neighboorhood $V'_l \subset W'_l$ of $-\beta_{j_l}$ such that $g_{l|V'_l} = 1$.
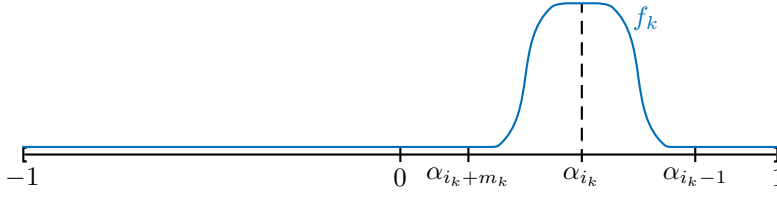
Fig. 4 Test functions for the proof of the geometric central limit theorem.

We fix an index k. Since $\frac{\lambda_i}{n} \to \alpha_i$ for all i, with probability going to 1 as n goes to infinity, the intersection of the support of X_λ with W_k consists exactly in

$$\left\{ \frac{\lambda_{i_k}}{n}, \frac{\lambda_{i_k+1}}{n}, \ldots, \frac{\lambda_{i_k+m_k-1}}{n} \right\},$$

that are even in V_k with big probability. So,

$$X_\lambda(f_k) = \frac{\lambda_{i_k} + \lambda_{i_k+1} + \cdots + \lambda_{i_k+m_k-1}}{n}$$

with probability going to 1; on the other hand, $X_\omega(f_k) = m_k\,\alpha_{i_k}$. Hence,

$$\Lambda_{i_k,m_k,n} = \nabla_{n,\omega}(f_k)$$

for any k with probability going to 1; similarly, $\Lambda'_{j_l,w_l,n} = \nabla_{n,\omega}(g_l)$ with big probability. Then, by Theorem 18, the random vector of the statement of our theorem converges in law to a centered gaussian vector

$$(\Lambda_{i_1,m_1,\infty}, \ldots, \Lambda_{i_r,m_r,\infty}, \Lambda'_{j_1,l_1,\infty}, \ldots, \Lambda'_{j_s,l_s,\infty}),$$

with $k(\Lambda_{i_k,m_k,\infty}, \Lambda_{i_l,m_l,\infty})$ equal to

$$X_\omega((xf_k(x))'(xf_l(x))') - X_\omega((xf_k(x))')\,X_\omega((xf_l(x))'),$$

and the same kind of formula with variables $\Lambda'_{j_k,w_k,\infty}$. Now, notice that $f'_k(x) = 0$ for any x in the support of X_ω, and any k. Hence, the previous covariance may be written as

$$X_\omega(f_k f_l) - X_\omega(f_k)\,X_\omega(f_l) = \delta_{kl}\,m_k\,\alpha_{i_k} - m_k m_l\,\alpha_{i_k}\alpha_{i_l},$$

and the proof for the other covariance formulas is the same with functions g_l instead of functions f_k. □

6 Pitman operators and a link with conditioned random walks and brownian matrices

In the previous paragraphs, we have shown that the rows and columns of a partition λ of law $\mathbb{P}_{n,\omega}$ have limit frequencies α_i and β_j, and gaussian fluctuations when gathered for equal coordinates of the point ω of the Thoma simplex. To treat the case of equal parameters, we shall use O'Connell's theory ([O'C03]) that relates the Robinson-Schensted algorithm to an operation on multidimensional paths, and therefore random tableaux coming from ω-shuffles to random walks conditioned to stay in a Weyl chamber. These random walks are also related to the processes of eigenvalues of brownian traceless Hermitian matrices. Thus, the fluctuations of a block of d rows or columns corresponding to equal coordinates of ω are the same as those of the eigenvalues of a gaussian matrix in $\mathfrak{su}(d,\mathbb{C})$, up to the additional gaussian term described by Theorem 19; see Theorem 25 for a precise statement. We shall prove this progressively, starting from the toy model of 2-shuffles, then introducing **Pitman operators** in order to treat the case of d-shuffles, and finally studying the general case.

6.1 The case of 2-shuffles

To begin with, let us treat minutely the case when

$$\omega = \left(\left(\frac{1}{2}, \frac{1}{2}, 0, 0, \dots \right), (0, 0, \dots) \right).$$

The corresponding ω-shuffles (which we shall call **2-shuffles**) are easy to describe. Indeed, for any n, a random permutation of law $\mathbb{Q}_{n,\omega}$ can be obtained by shuffling together two blocks of size n_1 and $n_2 = n - n_1$ with $n_1 \sim \mathcal{B}(n, \frac{1}{2})$, that is to say that

$$\mathbb{P}[n_1 = k] = \frac{1}{2^n} \binom{n}{k}.$$

We recall that all the $\binom{n}{k}$ shuffles of two blocks of sizes k and $n-k$ are equiprobable. Thus, if $b_1, \dots, b_n$ are independent identically distributed variables with $\mathbb{P}[b_i = 1] = \mathbb{P}[b_i = 2] = 1/2$, then one can produce a permutation of law $\mathbb{Q}_{n,\omega}$ by:

– setting $n_1 = \operatorname{card} \{i \mid b_i = 1\}$, $n_2 = \operatorname{card} \{i \mid b_i = 2\}$.
– sending the integers of $[\![1, n_1]\!]$ in this order to the positions i such that $b_i = 1$, and then sending the integers of $[\![n_1 + 1, n_1 + n_2]\!]$ in this order to the positions i such that $b_i = 2$.

For instance, if the sequence of Bernoulli variables is 121121221, then the corresponding permutation[7] will be 162374895. This explicit construction of 2-shuffles will enable us to relate them to random walks — here, the random walk that we shall consider is:

---

[7] With the vocabulary of §2, σ is the standardization of the word whose letters are the Bernoulli variables.
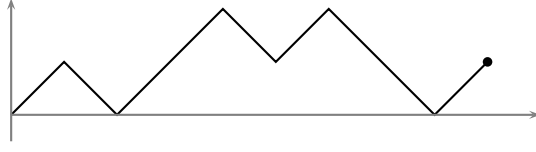
Fig. 5 Random walk associated to the 2-shuffle $162374895 = \mathrm{Std}(121121221)$.

Given a permutation σ obtained by standardization from a word w in $\{1,2\}^*$, the shape of the tableau $P(\sigma)$ is the same as the shape of the tableau $P(w)$. Actually, for any $k \le n$, if $P_k(\sigma)$ (respectively, $P_k(w)$) is the tableau obtained after Schensted insertion of the k first letters of σ (resp., of the k first letters of w), then $P_k(\sigma)$ has the same shape as $P_k(w)$. Hence, we will study the random process $(P(w_k))_{k \ge 0}$, where $w_k = b_1 b_2 \ldots b_k$ is a random word in $\{1,2\}^k$ with independent Bernoulli letters. This process takes its values in

$$\mathrm{SST}\{1,2\} = \{\text{semi-standard tableaux with entries in } \{1,2\}\}.$$

A semi-standard tableau with entries in $\{1,2\}$ has at most two rows, and the entries in the second row are all equal to 2. Thus, if one knows the size n of the tableau, then it is entirely determined by the number m_{11} of entries 1 on the first row, and by the number m_{12} of entries 2 on the first row:

with $m_{22} = n - m_{11} - m_{12}$. So, let us study in detail the random process $(m_{11}(k), m_{12}(k))$. For this purpose, we introduce the random walk $X_k = 3k - 2(b_1 + b_2 + \cdots + b_k)$; this is a standard random walk on $\mathbb{Z}$, starting from 0.

Lemma 20. *For any k, $m_{11}(k) = \frac{k+X_k}{2}$ and $m_{12}(k) = S_k - X_k$, where $S_k = \sup_{j \le k} X_j$.*

Proof. The quantity $m_{11}(k)$ is just the number of "1" in the sequence $b_1, \ldots, b_k$. Hence,

$$m_{11}(k) = \sum_{i=1}^k 2 - b_i = 2k - (b_1 + b_2 + \cdots + b_k) = \frac{k + X_k}{2}.$$

As for $m_{12}(k)$, let us prove the result by induction on k. Supposing $m_{12}(k) = S_k - X_k$, let us compute $m_{12}(k+1)$:

1. If $b_{k+1} = 1$, one inserts a letter 1 in the first row. Suppose that $m_{12}(k) = 0$. Then, $m_{12}(k+1)$ is also equal to 0, so $S_k = X_k$, and $X_{k+1} = X_k + 1$, because $b_{k+1} = 1$ means that we go up. This implies that $S_{k+1} = X_{k+1}$, so $m_{12}(k+1) = 0 = S_{k+1} - X_{k+1}$. On the contrary, suppose that $m_{12}(k) \geq 1$. Then, the letter 1 will replace a letter 2, so

$$m_{12}(k+1) = m_{12}(k) - 1 = S_k - X_k - 1.$$

   On the other hand, since $X_k = S_k - m_{12}(k) \leq S_k - 1$, $X_{k+1} = X_k + 1 \leq S_k$, so $S_k = S_{k+1}$. But then, $m_{12}(k+1) = S_k - (X_k + 1) = S_{k+1} - X_{k+1}$, so the identity holds again at rank $k+1$.

2. If $b_{k+1} = 2$, one inserts a letter 2 at the end of the first row, so

$$m_{12}(k+1) = m_{12}(k) + 1 = S_k - X_k + 1.$$

   On the other hand, $b_{k+1} = 2$ means that the random walk goes down, so $X_{k+1} = X_k - 1$ and $S_k = S_{k+1}$. Then, $m_{12}(k+1) = S_k - (X_k - 1) = S_{k+1} - X_{k+1}$.

Hence, if $\ell_k(\sigma)$ is the length of the longest increasing subword of $\sigma(1) \cdots \sigma(k)$, then

$$\ell_k(\sigma) = m_{11}(k) + m_{12}(k) = \frac{k}{2} + \frac{2S_k - X_k}{2},$$

and this for any $k \leq n$. $\square$

It is known since Pitman's works (see [Pit75]) that if $(X_k)_{k \in \mathbb{N}}$ is a standard random walk on $\mathbb{Z}$, then $(2S_k - X_k)_{k \in \mathbb{N}}$ is a markovian random process on the non-negative integers, with

$$p(k, k+1) = \frac{k+2}{2(k+1)} \qquad ; \qquad p(k, k-1) = \frac{k}{2(k+1)}.$$

After scaling by n on the x-axis and $\sqrt{n}$ on the y-axis, the random process $(2S_k - X_k)_{k \in [\![0,n]\!]}$ converges in law in $\mathscr{C}([0,1])$ to a **3-dimensional Bessel process** $(\mathrm{BES}^3_t)_{t \in [0,1]}$, whose trajectorial laws are the same as $(2S_t - B_t)_{t \in [0,1]}$, with $(B_t)_{t \in [0,1]}$ standard brownian motion and $S_t = \sup_{s \leq t} B_s$. Alternatively, $(\mathrm{BES}^3_t)_{t \in [0,1]}$ can be defined as the solution of the stochastic differential equation

$$dY_t = dB_t + \frac{dt}{Y_t},$$

see [RY91, Chapter 6, §3]. For this reason, and using Itô's calculus, one can show that $(\mathrm{BES}^3_t)_{t \in [0,1]}$ has the same trajectorial laws as the euclidian norm $(\sqrt{X_t^2 + Y_t^2 + Z_t^2})_{t \in [0,1]}$ of a 3-dimensional brownian motion. In particular, this allows to compute the density of the law of BES^3_t:

$$\mathrm{BES}^3_t \sim \mathbf{1}_{y \geq 0} \sqrt{\frac{2}{\pi \, t^3}} \, y^2 \, \mathrm{e}^{-\frac{y^2}{2t}} \, dy.$$

With these classical results from probability, Lemma 20 can be reinterpreted as follows:

Proposition 21. *Let σ be a random permutation obtained by a 2-shuffle. We denote*

$$L^{(n)}_{t \in [0,1[} = \frac{2}{\sqrt{n}} \left(\left[(\lfloor nt+1 \rfloor - nt)\, \ell_{\lfloor nt \rfloor}(\sigma) + (nt - \lfloor nt \rfloor)\, \ell_{\lfloor nt \rfloor + 1}(\sigma) \right] - \frac{nt}{2} \right)$$

the random process obtained by renormalization and interpolation of the process $(\ell_k(\sigma) - \frac{k}{2})_{k \in [\![0,n]\!]}$. As n goes to infinity, $(L^{(n)}_t)_{t \in [0,1]}$ converges in law[8] in $\mathscr{C}([0,1])$ to a 3-dimensional Bessel process $(\mathrm{BES}^3_t)_{t \in [0,1]}$ (see Figure 6).

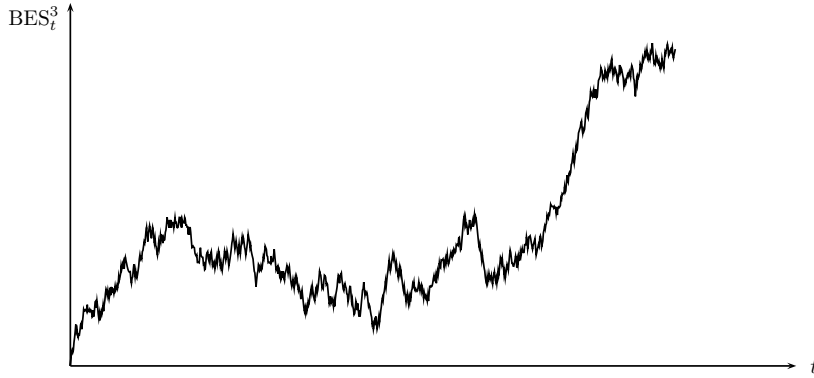**Fig. 6** A trajectory of the 3-dimensional Bessel process.

In particular,

$$\ell(\sigma) = \lambda_1 = \frac{n}{2} + \frac{\sqrt{n}}{2} Y + o\left(\sqrt{n}\right), \quad \text{with } Y \sim \mathrm{BES}^3_1 \sim \mathbf{1}_{y \geq 0} \sqrt{\frac{2}{\pi}}\, y^2\, e^{-\frac{y^2}{2}}\, dy.$$

To conclude this paragraph, let us give yet another representation of the 3-dimensional Bessel process. We denote by $\mathfrak{su}(2, \mathbb{C})$ the space of traceless hermitian matrices of size 2×2. An element of $\mathfrak{su}(2, \mathbb{C})$ can be written as

$$M = \begin{pmatrix} m_1 & m_2 + \mathrm{i}m_3 \\ m_2 - \mathrm{i}m_3 & -m_1 \end{pmatrix},$$

---

[8] As usual, the interpolation is only needed in order to work in $\mathscr{C}([0,1])$; if one uses instead step *càdlàg* functions, one has the same result of convergence, but this time in Skorohod's space $\mathscr{D}([0,1])$, see [Bil69, Chapter 3, §14].

and $\mathfrak{su}(2,\mathbb{C})$ is an euclidian space for the scalar product $\langle A \mid B \rangle = \operatorname{tr} AB^* = 2(a_1 b_1 + a_2 b_2 + a_3 b_3)$. A brownian motion in $\mathfrak{su}(2,\mathbb{C})$ is a continuous gaussian process $(M_t)_{t \in \mathbb{R}_+}$ starting from the null matrix, and with covariances

$$\mathbb{E}[\operatorname{tr}((AM_s)(BM_t)^*)] = K \,(s \wedge t) \operatorname{tr}(AB^*).$$

For instance, if X, Y, Z are three independent brownian motions in $\mathbb{R}$, then

$$\begin{pmatrix} X_t & Y_t + \mathrm{i}Z_t \\ Y_t - \mathrm{i}Z_t & -X_t \end{pmatrix}$$

is a brownian motion in $\mathfrak{su}(2,\mathbb{C})$, with $K = 3$. Now, the eigenvalues of such a matrix brownian motion are $\pm\sqrt{X_t^2 + Y_t^2 + Z_t^2}$, so they are described by a 3-dimensional Bessel process. Hence, Proposition 21 can be restated as follows: if $\lambda = (\lambda_1, \lambda_2)$ is a random partition of law $\mathbb{P}_{n,\omega}$ with $\omega = ((1/2, 1/2), (0))$, then

$$\frac{2}{\sqrt{n}} \left(\lambda_1 - \frac{n}{2}, \lambda_2 - \frac{n}{2} \right) \simeq_{n \to \infty} (x_1, x_2),$$

where $x_1 \geq x_2$ are the two eigenvalues of a brownian matrix in $\mathfrak{su}(2)$ at time $t = 1$. This is the good formulation of Proposition 21 for a generalization to the case of d-shuffles.

6.2 The case of d-shuffles

With a similar framework as before, we want to understand the asymptotics of **d-shuffles**, meaning that we consider the case when

$$\omega = \left(\left(\frac{1}{d}, \frac{1}{d}, \dots, \frac{1}{d}, 0, 0, \dots \right), (0) \right),$$

with d coordinates equal to $\frac{1}{d}$ in α. A generalization of the arguments of the previous paragraph to the case of d-shuffles has been proposed by N. O'Connell in [O'C03], in connection with previous results from [OY02]. Let us denote by Π_d the set of paths $W : \mathbb{N} \to \mathbb{N}^d$, such that $W(0) = (0, 0, \dots, 0)$ and

$$W(k+1) - W(k) \in \{0, e_1, e_2, \dots, e_d\}$$

for any k, where $e_i = (0_1, 0_2, \dots, 1_i, \dots, 0_d)$. If w is a word in $\{1, 2, \dots, d\}^*$ of length n, we associate to it a path $(W^{(k)})_{k \leq n}$ in Π_d, with

$$W(k+1) - W(k) = e_i \quad \text{if } w_k = i.$$

Now, for any pair of sequences $(x(k))_{k \in \mathbb{N}}$ and $(y(k))_{k \in \mathbb{N}}$ in Π_1, we define two new sequences

$$
\begin{aligned}
(x \curlywedge y)(k) &= \min\{x(j) - y(j) + y(k)\}_{j \leq k}; \\
(x \curlyvee y)(k) &= \max\{x(j) - y(j) + y(k)\}_{j \leq k}.
\end{aligned}
$$

Notice that the rules $\curlywedge$ and $\curlyvee$ are not associative. In the following, we will use the following convention: $x \curlywedge y \curlywedge z = (x \curlywedge y) \curlywedge z$ and $x \curlyvee y \curlyvee z = (x \curlyvee y) \curlyvee z$. We then define operators $D_d : \Pi_d \to \Pi_d$ and $T_d : \Pi_d \to \Pi_{d-1}$:

$$D_d(x_1, \ldots, x_d) = (x_1, x_1 \curlywedge x_2, x_1 \curlywedge x_2 \curlywedge x_3, \ldots, x_1 \curlywedge x_2 \curlywedge \cdots \curlywedge x_d);$$
$$T_d(x_1, \ldots, x_d) = (x_2 \curlyvee x_1, x_3 \curlyvee (x_1 \curlywedge x_2), \ldots, x_d \curlyvee (x_1 \curlywedge x_2 \curlywedge \cdots \curlywedge x_{d-1})).$$

Finally, from a path W in Π_d, we define new paths $D^i W \in \Pi_{d-i+1}$ and $T^i W \in \Pi_{d-i}$ by the recursive rule

$$D^1 W = D_d(W_1, W_2, \ldots, W_d) \quad ; \quad T^1 W = T_d(W_1, \ldots, W_d)$$
$$D^{i+1} W = D_{d-i}(T^i W) \quad ; \quad T^{i+1} W = T_{d-i}(T^i W).$$

For any path W, the family $(D^1 W, D^2 W, \ldots, D^d W)$ is a sequence of triangular arrays. The link between this construction and the RSK algorithm is given by the following result of O'Connell:

Theorem 22. *For any word $w \in \{1, 2, \ldots, d\}^n$, the semi-standard tableau of size n obtained from w by Schensted insertion is*

m_{22} "2"	m_{23} "3"	$\cdots$	m_{2d} "d"	
m_{11} "1"	m_{12} "2"		$\cdots$	m_{1d} "d"

where[9] $m_{ij} = (D^{j-i+1}W)_i(n) - (D^{j-i}W)_i(n)$ for any $1 \le i \le j \le d$, W being the path associated to the word w. In particular, for any $k \le n$, the shape $(\lambda_1, \lambda_2, \ldots, \lambda_d)$ of $P_k(w)$ is

$$\left((D^d W)_1(k), (D^{d-1}W)_2(k), \ldots, (D^1 W)_d(k)\right).$$

We refer to [O'C03, Theorem 3.1] for a proof of this important result. Now, we have to understand the transformation

$$G_d : \Pi_d \to \Pi_d$$
$$W \mapsto \left((D^d W)_1, (D^{d-1}W)_2, \ldots, (D^1 W)_d\right),$$

and what is the law of $G_d(W)$ when W is a random path in $\mathbb{N}^d$ associated to a random word in $\{1, 2, \ldots, d\}^*$. To begin with, we shall notice that the operators $\curlywedge$ and $\curlyvee$ also make sense for any continuous or càdlàg functions $x(t)$ and $y(t)$. Thus, we can define operators

$$D_d : \mathscr{C}_0(\mathbb{R}_+, \mathbb{R}^d) \to \mathscr{C}_0(\mathbb{R}_+, \mathbb{R}^d)$$
$$T_d : \mathscr{C}_0(\mathbb{R}_+, \mathbb{R}^d) \to \mathscr{C}_0(\mathbb{R}_+, \mathbb{R}^{d-1})$$
$$G_d : \mathscr{C}_0(\mathbb{R}_+, \mathbb{R}^d) \to \mathscr{C}_0(\mathbb{R}_+, \mathbb{R}^d)$$

---

[9] By convention, $D^0 W = (0, 0, \ldots, 0)$ for any path W, so $m_{ii} = (D^1 W)_i(n)$ for any i.

where $\mathscr{C}_0(\mathbb{R}_+, \mathbb{R}^d)$ denotes the set of continuous functions $w(t)$ from $\mathbb{R}_+$ to $\mathbb{R}^d$, with $w(0) = (0, 0, \ldots, 0)$. Moreover, G_d has the following property, see [O'C03, Lemma 2.2]: for any path $w \in \mathscr{C}_0(\mathbb{R}_+, \mathbb{R}^d)$ and any time $t \in \mathbb{R}_+$,

$$w_1(t) + w_2(t) + \cdots + w_d(t) = (G_d w)_1(t) + (G_d w)_2(t) + \cdots + (G_d w)_d(t).$$

In particular, G_d stabilizes the subspace $\mathscr{C}_0(\mathbb{R}_+, H) \subset \mathscr{C}_0(\mathbb{R}_+, \mathbb{R}^d)$ of functions taking their values in the hyperplane

$$H = \{(x_1, x_2, \ldots, x_d) \mid x_1 + x_2 + \cdots + x_d = 0\}.$$

On the other hand, for any x, y, f functions in $\mathscr{C}_0(\mathbb{R}_+, \mathbb{R})$, one has

$$(x + f) \curlywedge (y + f) = (x \curlywedge y) + f \qquad ; \qquad (x + f) \curlyvee (y + f) = (x \curlyvee y) + f.$$

This implies immediately that:

$$G_d(x_1 + f, x_2 + f, \ldots, x_d + f) = G_d(x_1, x_2, \ldots, x_d) + (f, f, \ldots, f).$$

In particular, set $f = \frac{x_1 + x_2 + \cdots + x_d}{d}$. Then,

$$G_d(x_1, x_2, \ldots, x_d) = (f, f, \ldots, f) + G_d(x_1 - f, x_2 - f, \ldots, x_d - f),$$

and the last term is a function in $\mathscr{C}_0(\mathbb{R}_+, H)$; hence, it suffices to study the restriction of G_d to the subspace $\mathscr{C}_0(\mathbb{R}_+, H)$.

The operators G_d are better understood in terms of the **Pitman transforms** attached to a root system (*cf.* [BBO05]). The hyperplane $H \subset \mathbb{R}^d$ contains a root system of type A_{d-1}, whose simple roots are $e_1 - e_2, e_2 - e_3, \ldots, e_{d-1} - e_d$. To each simple root $\delta_i = e_i - e_{i+1}$, we attach an operator $P_i : \mathscr{C}_0(\mathbb{R}_+, \mathbb{R}^d) \to \mathscr{C}_0(\mathbb{R}_+, \mathbb{R}^d)$, defined by the following formula:

$$(P_i w)(t) = w(t) - (\min\{\langle \delta_i \mid w(s) \rangle\}_{s \leq t}) \, \delta_i.$$

Notice that each P_i stabilizes the subspace $\mathscr{C}_0(\mathbb{R}_+, H)$. In [BBO05], it is shown that the operators P_i satisfy the braid relations of the generators $s_i = (i, i-1)$ of the Coxeter group $\mathfrak{S}_d$:

$$\forall i \leq d - 2, \ \ P_i P_{i+1} P_i = P_{i+1} P_i P_{i+1};$$
$$\forall i, j \leq d - 1, \ \ |i - j| \geq 2 \Rightarrow P_i P_j = P_j P_i.$$

It enables us to define P_σ for any $\sigma \in \mathfrak{S}_d$, and it is not hard to see that $P_\sigma|_{\mathscr{C}_0(\mathbb{R}_+, H)}$ takes its values in

$$\mathscr{C}_0\left(\mathbb{R}_+, \bigcap_{s_i \in L_\sigma} H_i^+\right),$$

where $L_\sigma = \{s_i \mid \ell(s_i \sigma) < \ell(\sigma)\}$, and $H_i^+ = \{x \in H \mid \langle \delta_i \mid x \rangle \geq 0\}$ — *cf.* [BBO05, Proposition 2.8]. In particular, if ω_0 is the permutation of maximal

length in $\mathfrak{S}_d$, that is to say $(d, d-1, \ldots, 1)$, then $P_{\omega_0}|_{\mathscr{C}_0(\mathbb{R}_+, H)}$ is an idempotent operator taking its values in $\mathscr{C}_0(\mathbb{R}_+, C)$, where $C = \bigcap_{i=1}^{d-1} H_i^+$ is the closed Weyl chamber in H corresponding to the simple roots. In terms of coordinates, notice that $x = (x_1, \ldots, x_d)$ is in C if and only if

$$x_1 + \cdots + x_d = 0 \quad \text{and} \quad x_1 \geq x_2 \geq \cdots \geq x_d.$$

Now, if ϕ_k^d is the projection $(x_1, \ldots, x_d) \in \mathbb{R}^d \mapsto (x_1, \ldots, x_k) \in \mathbb{R}^k$, then it is easily seen that

$$T_d(W) = \phi_{d-1}^d \, P_{d-1} P_{d-2} \cdots P_2 P_1(W);$$
$$T^i W = \phi_{d-i}^d \, (P_{d-i} \cdots P_1) \, (P_{d-i+1} \cdots P_1) \, \cdots \, (P_{d-1} \cdots P_1)(W).$$

From this, one sees that $G_d = P_1(P_2 P_1) \cdots (P_{d-1} \cdots P_1) = P_{\omega_0}$. In particular, $G_d|_{\mathscr{C}_0(\mathbb{R}_+, H)}$ takes its values in $\mathscr{C}_0(\mathbb{R}_+, C)$, which was not clear from the first definition.

Then, let us consider the path W associated to a random word $w \in \{1, 2, \ldots, d\}^*$, where each letter of w is independent of the other letters, and uniformly distributed on $[\![1, d]\!]$. If

$$\widetilde{W}(k) = W(k) - \left(\frac{k}{d}, \frac{k}{d}, \ldots, \frac{k}{d} \right),$$

then $\widetilde{W}$ lives in the hyperplane H, and by (multi-dimensional) Donsker's theorem (*cf.* [Bil69, Chapter 2]), after time-space scaling, $\widetilde{W}$ converges to a brownian motion $(w_t)_{t \in \mathbb{R}}$ in H, that is to say a continuous gaussian process in H with covariances

$$\mathbb{E}[\langle a \mid w_s \rangle \, \langle b \mid w_t \rangle] = \frac{1}{d} \, (s \wedge t) \, \langle a \mid b \rangle \, .$$

Indeed, if X is a random step $\widetilde{W}(n+1) - \widetilde{W}(n)$, then $\langle \delta_i \mid X \rangle$ equals 1 if $X = e_i$, -1 if $X = e_{i+1}$, and 0 otherwise. From this, one sees that $\mathbb{E}[\langle \delta_i \mid X \rangle \, \langle \delta_j \mid X \rangle]$ equals $\frac{2}{d}$ if $i = j$, $-\frac{1}{d}$ if $i = j \pm 1$, and 0 otherwise. But this is exactly $\frac{\langle \delta_i \mid \delta_j \rangle}{d}$, whence the result after applying Donsker's theorem.

Let us forget for the moment the factor $\frac{1}{d}$, and consider a standard brownian motion w in H. Then, it can be shown that $P_{\omega_0}(w)$ is a brownian motion conditioned in Doob's sense to stay in the Weyl chamber C, *cf.* [BBO05, Theorem 5.6] or [Bia09]. We have drawn on Figure 7 a random walk in $H \subset \mathbb{R}^3$ corresponding to a random word (in blue), and its image by Pitman's transform $P_{\omega_0} = P_1 P_2 P_1$ (in violet); it approximates a brownian motion conditioned to stay in a Weyl chamber[10].

---

[10] On the drawing, it appears that an important part of the trajectory of the initial brownian motion is simply translated by Pitman's transform P_{ω_0}. This observation can be made precise, and one obtains a generalization of Williams' decomposition of the trajectories of the BES^3 process, *cf.* [RY91, Chapter 6, Theorem 3.11].
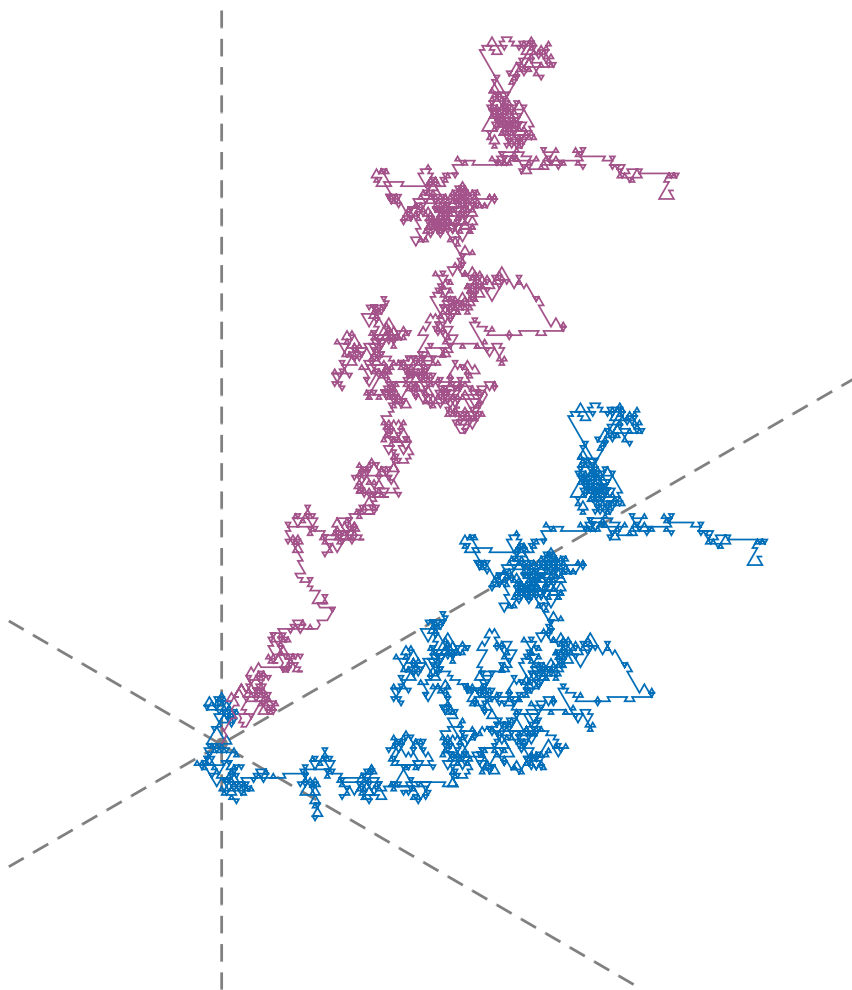
Fig. 7 Action of Pitman's transform on a brownian motion when $d = 3$.

This conditioned brownian motion is the Markov process on C with transition probabilities

$$q_t(x, y) = \frac{\Delta(y)}{\Delta(x)} \sum_{\sigma \in \mathfrak{S}_d} \varepsilon(\sigma) \, p_t(x, \sigma(y))$$

where $p_t(x, y) = (2\pi t)^{-\frac{d-1}{2}} \exp\left(-\frac{\|x-y\|^2}{2t}\right)$ is the transition kernel of the standard brownian motion in H, and

$$\Delta(x) = \prod_{\alpha \text{ positive root}} \langle \alpha \mid x \rangle = \prod_{1 \leq i < j \leq d} x_i - x_j$$

is (up to a multiplicative constant) the unique positive function on C that is harmonic with respect to the symmetrized kernel and that vanishes at the boundary of C. The singularity of the transition kernel at $x = (0, 0, \ldots, 0)$ is solvable[11], with

$$q_t(0, y) = \frac{\Delta(y)^2}{1! \, 2! \, \cdots \, d - 1! \, t^{\frac{d(d-1)}{2}}} \, p_t(0, y).$$

This Markov process corresponds to the process of the eigenvalues

$$(x_{t,1} \geq x_{t,2} \geq \cdots \geq x_{t,d})$$

of a brownian matrix in $\mathfrak{su}(d, \mathbb{C})$ with covariances

$$\mathbb{E}[\mathrm{tr}((AM_s)(BM_t)^*)] = \frac{d^2 - 1}{d} \, (s \wedge t) \, \mathrm{tr}(AB^*),$$

see [Dys62, Bia09], and [AGZ09, Theorem 2.5.2] for checking that the multiplicative constant is indeed $\frac{d^2-1}{d}$. Taking account of the factor $\frac{1}{d}$ put aside before, we get:

Theorem 23. *In the following, $(M_t)_{t \in \mathbb{R}_+}$ is a brownian matrix in $\mathfrak{su}(d, \mathbb{C})$ starting from the null matrix and with normalized covariances*

$$\mathbb{E}[\mathrm{tr}((AM_s)(BM_t)^*)] = \frac{d^2 - 1}{d^2} \, (s \wedge t) \, \mathrm{tr}(AB^*);$$

and $(x_{1,t} \geq x_{2,t} \geq \cdots \geq x_{d,t})$ is the set of eigenvalues of M_t. On the other hand, we consider a random permutation σ obtained by a d-shuffle, and we denote

$$L_{i,t}^{(n)} = \frac{1}{\sqrt{n}} \left(\left[(\lfloor nt + 1 \rfloor - nt) \, \Lambda_{i,\lfloor nt \rfloor}(\sigma) + (nt - \lfloor nt \rfloor) \, \Lambda_{i,\lfloor nt \rfloor + 1}(\sigma) \right] - \frac{nt}{d} \right),$$

where $(\Lambda_{1,k}(\sigma), \ldots, \Lambda_{d,k}(\sigma))$ is the shape of the tableau $P_k(\sigma)$ obtained by Schensted insertion from the k first letters of σ. As n goes to infinity,

$$\left(L_{1,t}^{(n)}, L_{2,t}^{(n)}, \ldots, L_{d,t}^{(n)} \right)_{t \in [0,1]} \to (x_{1,t}, x_{2,t}, \ldots, x_{d,t})_{t \in [0,1]},$$

where the convergence is in law and takes place in $\mathscr{C}([0,1], \mathbb{R}^d)$. In particular, if $\lambda \sim \mathbb{P}_{n,\omega}$ with $\omega = ((1/d, 1/d, \ldots, 1/d), (0))$, then[12]

$$\lambda_i = \frac{n}{d} + \sqrt{\frac{n}{d}} \, Y_i + o\left(\sqrt{n}\right)$$

with

$$(Y_1, \ldots, Y_d) \sim \mathbf{1}_{\substack{y_1 \geq y_2 \geq \cdots \geq y_d \\ y_1 + y_2 + \cdots + y_d = 0}} \frac{\Delta(y_1, y_2, \ldots, y_d)^2}{1! \, 2! \, \cdots \, d - 1! \, (2\pi)^{\frac{d-1}{2}}} \, \mathrm{e}^{-\frac{\|y\|^2}{2}} \, dy.$$

---

[11] This is easily seen by expanding terms $\exp(-\|\eta_i - x_{\sigma(i)}\|^2/2t)$ with $\eta \to 0$, and by using the theory of antisymmetric polynomials, see [Mac95, §1.3].

[12] Here, notice that dy denotes the $(d-1)$-dimensional Lebesgue measure on H.

6.3 The general case of equal coordinates

We pursue the study of the fluctuations of rows of a tableau obtained by Schensted insertion from a generalized shuffle, and we consider now the case when $\omega = ((p_1^{d_1}, p_2^{d_2}, \ldots, p_r^{d_r}), (0))$ with

$$p_1 > p_2 > \cdots > p_r > 0 \quad \text{and} \quad d_1 p_1 + d_2 p_2 + \cdots + d_r p_r = 1.$$

Here, by $p_i^{d_i}$, we mean that p_i appears d_i times in α. It will be clear from our reasoning that one could even consider an infinite sequence of p_i's with the same property. So, if one understands the behaviour of rows with a finite number of parameters, the study of the general case when $\omega = (\alpha, 0)$ and $\gamma = 0$ will be completed. In the following, we set $d = d_1 + d_2 + \cdots + d_r$.

As before, an ω-shuffle of size n corresponding to parameters $p_1^{d_1}, \ldots, p_r^{d_r}$ can be obtained by standardizing a random word $w = b_1 b_2 \ldots b_n$ in $[\![1, d]\!]^n$, where the letters b_k are independent and of law

$$\mathbb{P}[b_k = i] = p_j \quad \text{if } d_1 + \cdots + d_{j-1} + 1 \leq i \leq d_1 + \cdots + d_j.$$

To such a (random) word, we attach as before a random walk $W \in \mathbb{Z}^d$, with $W(k+1) - W(k) = e_i$ if $b_k = i$. O'Connell's Theorem 22 holds again, so $G_d(W)(k)$ is the form $\Lambda_k(\sigma)$ of the tableau $P_k(\sigma)$ (or $P_k(w)$) obtained by Schensted insertion from the k first letters of σ. Now, we shall use the following technical lemma:

Lemma 24. *If $W = (x_1, \ldots, x_d)$ is a random walk obtained from a shuffle of parameter $p_1^{d_1}, \ldots, p_r^{d_r}$, we introduce random functions $f_1, \ldots, f_d$ such that $(f_1, \ldots, f_d)$ is the vector*

$$G_d(x_1, \ldots, x_d) - (G_{d_1}(x_1, \ldots, x_{d_1}), \ldots, G_{d_r}(x_{d_1 + \cdots + d_{r-1} + 1}, \ldots, x_{d_1 + \cdots + d_r})).$$

These functions are almost surely bounded (as functions of $\mathbb{N}$).

Example. Before proving Lemma 24, let us see study the simplest example where it applies, namely, when $\omega = ((a, 1 - a), (0))$ with $a > 1/2$. The random walk to consider is two dimensional and will be denoted $w_n = (x_n, y_n)_{n \in \mathbb{N}}$; moreover,

$$(G_2 w)_n = (x_n - \min\{x_k - y_k\}_{k \leq n}, y_n + \min\{x_k - y_k\}_{k \leq n}).$$

However, $(z_n = x_n - y_n)_{n \in \mathbb{N}}$ is a random walk on $\mathbb{Z}$ with $\mathbb{P}[z_{n+1} = z_n + 1] = a$, $\mathbb{P}[z_{n+1} = z_n - 1] = 1 - a$, and $z_0 = 0$ almost surely. It is well-known that such a random walk is transient, with

$$\mathbb{P}\left[\inf_{n \in \mathbb{N}} z_n \leq -m\right] = \left(\frac{1 - a}{a}\right)^m$$

$$\mathbb{P}\left[\min_{n \in \mathbb{N}} z_n = -m\right] = \left(\frac{1 - a}{a}\right)^m - \left(\frac{1 - a}{a}\right)^{m+1} = \frac{2a - 1}{a}\left(\frac{1 - a}{a}\right)^m.$$

Thus, $(G_2(x, y)) - (x, y) = (-z, z)$, with $|z|$ almost surely bounded by a (random) constant of expectation $\frac{1 - a}{2a - 1}$.

Proof of Lemma 24. The example above can be generalized by using the language of Pitman operators: if P_i acts on the random walk with $i = d_1 + \cdots + d_j$ for some $j \in [\![1, r-1]\!]$, then $P_i w$ differs from w by almost surely bounded functions on the coordinates of index $d_1 + \cdots + d_j$ and $d_1 + \cdots + d_j + 1$. This is exactly for the same reason: the difference of the two random walks $x_{d_1 + \cdots + d_j}$ and $x_{d_1 + \cdots + d_j + 1}$ is almost surely bounded from below. Quite obviously, this is also true if one replaces w by $w + \delta$ with δ almost surely bounded vector, so one can use this argument several times. More generally, this works also if w is replaced by $P_\sigma(w)$, where σ is a permutation in the Young subgroup $\mathfrak{S}_{d_1} \times \mathfrak{S}_{d_2} \times \cdots \times \mathfrak{S}_{d_r} \subset \mathfrak{S}_d$; indeed, such an operator won't change the fact that the coordinates $x_{d_1 + \cdots + d_j}$ and $x_{d_1 + \cdots + d_j + 1}$ have different rates of growth to infinity.

In the operator $G_d = P_{\omega_0} = P_1 (P_2 P_1)(P_3 P_2 P_1) \cdots (P_{d-1} \cdots P_1)$, it allows to neglict all the factors P_i with $i = d_1 + \cdots + d_j$: they will simply modify the random path by almost surely bounded functions. So, $G_d w \simeq J_d w$, where J_d is obtained from G_d by deleting in the reduced decomposition of ω_0 all the s_i's with $i = d_1 + \cdots + d_j$, $j \in [\![1, r-1]\!]$. Then, J_d acts separately on each block of coordinates $x_{d_1 + \cdots + d_{j-1} + 1}, \ldots, x_{d_1 + \cdots + d_j}$, and the action corresponding to the block of size d_j is given by the operator

$$G_{d_j} \left(P_{d_j - 1} \cdots P_2 P_1 \right)^k,$$

where $k = d - d_j$. Now, each P_i is an idempotent operator (this is almost obvious from the definition given on page 48), and on the other hand, for any i, there is a reduced decomposition of ω_0 in $\mathfrak{S}_{d_j}$ that ends by s_i. Hence,

$$G_{d_j} \left(P_{d_j - 1} \cdots P_2 P_1 \right)^k = G_{d_j},$$

whence the result. $\qquad\square$

Now, for each block $x_{d_1 + \cdots + d_{j-1} + 1}, \ldots, x_{d_1 + \cdots + d_j}$, one can follow the same reasoning as in the case of d-shuffles, with some precaution. First, notice that if $g_j = \frac{1}{d_j} \sum_{i=1}^{d_j} x_{d_1 + \cdots + d_{j-1} + i}$, then:

1. In $\mathbb{N}^r$, $(d_1 g_1, \ldots, d_r g_r)$ is a random walk of parameter $d_1 p_1, \ldots, d_r p_r$.

2. The vectors

$$\widetilde{W}_j = (x_{d_1 + \cdots + d_{j-1} + 1} - g_j, \ldots, x_{d_1 + \cdots + d_{j-1} + d_j} - g_j)$$

with $j \in [\![1, r]\!]$ are random walks in $H \subset \mathbb{R}^{d_j}$, and they are all independent, and independent from the vector $(d_1 g_1, \ldots, d_r g_r)$.

One takes account of the first vector by the standard central limit theorem:

$$d_i g_i = n(d_i p_i) + \sqrt{n}\, X_i + o\left(\sqrt{n}\right)$$

where $(X_1, \ldots, X_r)$ is a gaussian vector of covariance matrix $k(X_i, X_j) = \delta_{ij}\, d_i p_i - d_i d_j\, p_i p_j$. More precisely, one can describe asymptotically

$$(d_1 g_1(k) - k\, d_1 p_1, \ldots, d_r g_r(k) - k\, d_r p_r)_{k \le n}$$

by using the multidimensional Donsker theorem. As for the vectors $\widetilde{W}_j$, they can be studied with the framework of the previous paragraph on d-shuffles. The only difference is the multiplicative constant in the covariances of the brownian motion that is the limit of $\widetilde{W}_j$ after time-space scaling: it is p_j instead of $\frac{1}{d_j}$. In the end, one obtains the following:

Theorem 25. *With the same notations as before, we fix independent brownian motions $(M_{t,j})_{t \in \mathbb{R}_+, 1 \le j \le r}$ taking values in the spaces $\mathfrak{su}(d_j, \mathbb{C})$, with covariances*

$$\mathbb{E}[\mathrm{tr}((AM_{s,j})(BM_{t,j})^*)] = \frac{d_j^2 - 1}{d_j}\, p_j\, (s \wedge t)\, \mathrm{tr}(AB^*);$$

$(x_{1,j,t} \ge x_{2,j,t} \ge \cdots \ge x_{d_j,j,t})$ will denote the set of eigenvalues of $M_{t,j}$. We also consider an independent brownian motion $(B_{1,t}, \ldots, B_{r,t})$ in $\mathbb{R}^d$, with covariances

$$\mathbb{E}[\langle X \mid B_s \rangle \, \langle Y \mid B_t \rangle] = (s \wedge t)\, X^t Q Y, \quad \text{where } Q = \left(\delta_{ij}\, \frac{p_i}{d_i} - p_i p_j \right)_{1 \le i,j \le r}.$$

On the other hand, we consider a random permutation obtained by a shuffle of parameters $p_1^{d_1}, \ldots, p_r^{m_r}$, and we denote as before $\Lambda_k(\sigma)$ the shape of the standard tableau $P_k(\sigma)$, and

$$L_{i,t}^{(n)} = \frac{1}{\sqrt{n}} \left(\left[(\lfloor nt + 1 \rfloor - nt)\, \Lambda_{i, \lfloor nt \rfloor}(\sigma) + (nt - \lfloor nt \rfloor)\, \Lambda_{i, \lfloor nt \rfloor + 1}(\sigma) \right] - p_j nt \right),$$

assuming that $d_1 + \cdots + d_{j-1} + 1 \le i \le d_1 + \cdots + d_j$. As n goes to infinity, $\left(L_{1,t}^{(n)}, \ldots, L_{i,t}^{(n)}, \ldots, L_{d,t}^{(n)} \right)_{t \in [0,1]}$ converges in law to

$$(B_{1,t} + x_{1,1,t}, \ldots, B_{j,t} + x_{i,j,t}, \ldots, B_{r,t} + x_{d_r,r,t})_{t \in [0,1]},$$

in $\mathscr{C}([0,1], \mathbb{R}^d)$. In particular, if $\lambda = \Lambda_n(\sigma) = \Lambda(\sigma)$ with σ chosen randomly as before, then

$$\lambda_i = n p_j + \sqrt{n p_j}\, (X_j + Y_{i,j}) + o\left(\sqrt{n}\right),$$

with

$$(Y_{1,j}, \ldots, Y_{d_j,j}) \sim \mathbf{1}_{\substack{y_{1,j} \ge y_{2,j} \ge \cdots \ge y_{d_j,j} \\ y_{1,j} + y_{2,j} + \cdots + y_{d_j,j} = 0}} \frac{\Delta(y_{1,j}, y_{2,j}, \ldots, y_{d_j,j})^2}{1!\, 2! \, \cdots \, d_j - 1!\, (2\pi)^{\frac{d_j - 1}{2}}}\, \mathrm{e}^{-\frac{\|y_j\|^2}{2}}\, dy_j$$

and $(X_1, \ldots, X_r)$ gaussian vector of covariance matrix

$$Q' = \left(\frac{\delta_{ij}}{d_i} - \sqrt{p_i p_j} \right)_{1 \le i,j \le r}.$$

Remark. At the end of Section 4, we have seen that the law of large numbers can be informally restated by saying that the longest increasing subsequences in a α-shuffle are not bigger at order n than the lengths of the blocks shuffled together. Theorem 25 ensures that it is also true at order $\sqrt{n}$, unless some blocks have the same frequency $p_j = \alpha_i = \alpha_{i'}$. Then, some of the increasing subwords become (stochastically) bigger than the sizes of the blocks to be shuffled, because of the terms $\sqrt{np_j}\, Y_{i,j}$.

Theorem 25 seems to remain true when $\beta \neq 0, \gamma \neq 0$, and in this case, the fluctuations of the first renormalized columns seem to be given by analogous formulas as those given above, with fluctuations "of type X" described by Theorem 19, and fluctuations "of type Y" (for equal coordinates) independent of the fluctuations of the rows. We don't know how to prove this, but we can give arguments in favor of this conjecture. These arguments can be made rigorous in some cases, for instance when $\gamma = 0$ and β has only a finite number of non-zero coordinates. When performing the Schensted insertion from a permutation σ obtained by an ω-shuffle, let us mark by a point • the boxes containing a number coming from a block of size m_j, and by a star ⋆ the boxes containing a number coming from the block of size l.

Example. Suppose $n = 16$, $n_1 = n_2 = 3$, $m_1 = m_2 = 3$, $l = 4$. We denote the "cards" to be shuffled by `0,1,...,9,A,B,...,F`. After cutting the deck, the five blocks to be shuffled are
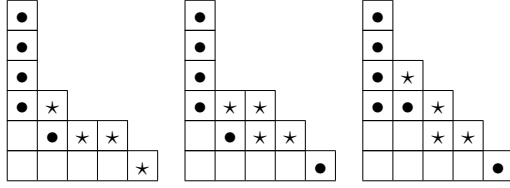
<div align="center">

`012, 345, 678, 9AB, CDEF`.

</div>

The blocks `678` and `9AB` are reversed, and `CDEF` is randomized, becoming from instance `DFCE`. Then, all the blocks are shuffled together, and one can for instance obtain
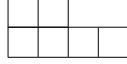
<div align="center">

`8B0D3A97FC41E562`.

</div>

Now, the partitions appearing in the RSK algorithm are:
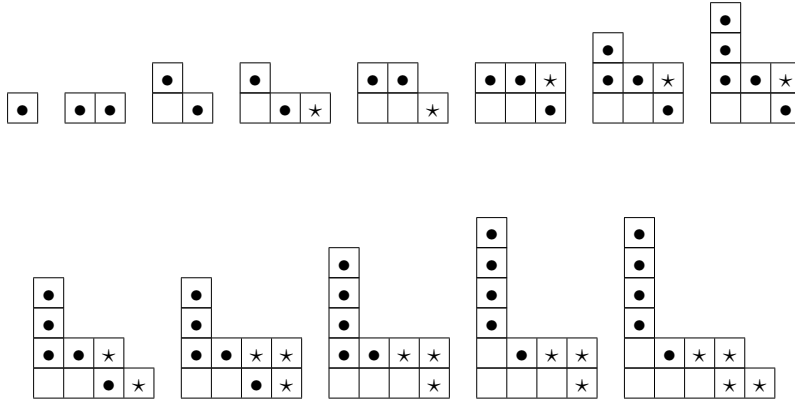
Notice that the partition of unmarked boxes

is the one corresponding to 034152; more generally, during the Schensted insertion, the partition of unmarked boxes evolves as the partition obtained *via* RSK from the subword containing only the letters coming from the blocks of sizes n_i. Thus, to prove that Theorem 25 holds again when $\beta \neq 0, \gamma \neq 0$ amounts to control the number of marked boxes appearing at the right of the first rows, and to show that it is a $o(\sqrt{n})$ — then, by symmetry of the problem, we would obtain the fluctuations of the columns, and see that they are analogous to those of the rows.

One should notice that each time a unmarked box appears in a row, it pushes a marked box

$$\boxed{\bullet} \quad \text{or} \quad \boxed{\star}$$
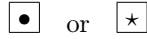
to the next row. Moreover, marked boxes push themselves to upper rows: for instance, if $\gamma = 0$ and $\beta = (\beta_1, \ldots, \beta_d)$, then one can not find more than d marked boxes on the same row. So the idea would be to show that marked boxes are "sufficiently pushed up". Unfortunately, to do this rigorously, it seems that we would need precise estimates on the probabilities of creating a "long" increasing subword in a shuffle of parameter $\alpha = 0$; and we don't have for the moment such estimates.

To conclude this section, let us look again at the partition-valued random process $(\Lambda_k(\sigma))_{k \in [\![0,n]\!]}$, where σ is obtained from an ω-shuffle, and $\Lambda_k(\sigma)$ is the shape of the tableau obtained by Robinson-Schensted insertion from the k first letters of σ. We have shown previously that after time-space scaling, and at least when $\beta = \gamma = 0$, the coordinates of this process converge jointly to a markovian process that can be written in terms of (conditioned) brownian motions. It raises the following question: is $(\Lambda_k(\sigma))_{k \in [\![0,n]\!]}$ itself a markovian process in $\mathfrak{Y} = \bigsqcup_{n \in \mathbb{N}} \mathfrak{Y}_n$? The answer is yes, and more precisely:

Theorem 26. *The random process $(\Lambda_k(\sigma))_{k \in [\![0,n]\!]}$ is markovian, with transition probabilities equal to*

$$p(\lambda, \Lambda) = \mathbf{1}_{\lambda \nearrow \Lambda} \, \frac{s_\Lambda(\omega)}{s_\lambda(\omega)}.$$

Proof. This is a generalization of [O'C03, Theorem 6.1], and the proof is exactly the same. The statement of Theorem 26 amounts to ask that for any sequence of partitions $\lambda^{(1)} \nearrow \cdots \nearrow \lambda^{(k)}$,

$$\mathbb{P}[\Lambda_1(\sigma) = \lambda^{(1)}, \cdots, \Lambda_k(\sigma) = \lambda^{(k)}] = s_{\lambda^{(k)}}(\omega).$$

However, one can expand $s_{\lambda^{(k)}}(\omega)$ by using the structure of Hopf algebra of **Sym**:

$$s_{\lambda^{(k)}}(\omega) = \big((1 \otimes \upsilon \otimes 1) \circ \Delta^3\big) s_{\lambda^{(k)}}(\alpha, \beta, \gamma E)$$
$$= \sum_{\nu \subset \mu \subset \lambda^{(k)}} s_\nu(\alpha)\, s_{\mu'\backslash\nu'}(\beta)\, s_{\lambda^{(k)}\backslash\mu}(\gamma E).$$

Then, the arguments are exactly the same as in [O'C03]. $\qquad\square$

7 Jones-Ocneanu traces of Hecke algebras and (q,t)-Plancherel measures

In this last section, we return to the "algebraic" study of the probability measures $\mathbb{P}_{n,\omega}$, and we introduce new observables in $\mathscr{O}$ related to the representation theory of the Hecke algebras $\mathscr{H}(\mathfrak{S}_n)$, that are deformations of the symmetric group algebras $\mathbb{C}\mathfrak{S}_n$. We then determine their asymptotic behaviour under the probability measures $\mathbb{P}_{n,\omega}$, thereby proving a q-analog of Theorem 17. Finally, we study the systems of coherent measures coming from the Jones-Ocneanu traces of the Hecke algebras (*cf.* §7.2); they form a (q,t)-deformation of the system of Plancherel measures described on page 11, see §7.3.

7.1 Hecke algebras of the symmetric groups

Let us fix a non-negative integer n. We recall that the (generic) **Iwahori-Hecke algebra** of the symmetric group $\mathfrak{S}_n$ is the $\mathbb{C}(q)$-algebra $\mathscr{H}(\mathfrak{S}_n)$ generated by elements $T_1, T_2, \ldots, T_{n-1}$, with presentation:

$$\begin{cases} \forall i \leq n-1, \ (T_i - q)(T_i + 1) = 0\,; \\ \forall i \leq n-2, \ T_i T_{i+1} T_i = T_{i+1} T_i T_{i+1}\,; \\ \forall i, j, \ |j-i| \geq 2 \Rightarrow T_i T_j = T_j T_i. \end{cases}$$

A linear basis of $\mathscr{H}(\mathfrak{S}_n)$ consists in the $T_\sigma = T_{i_1} T_{i_2} \cdots T_{i_{\ell(\sigma)}}$, where σ runs over $\mathfrak{S}_n$ and $\sigma = s_{i_1} s_{i_2} \cdots s_{i_{\ell(\sigma)}}$ is any reduced decomposition of σ in elementary transpositions. Thus, $\mathscr{H}(\mathfrak{S}_n)$ is a quantization of the symmetric group algebra $\mathbb{C}\mathfrak{S}_n$. In fact, it has the same representation theory: it is a semisimple $\mathbb{C}(q)$-algebra, and its irreducible modules $V^\lambda(q)$ are labelled by the integer partitions of size n and have the same dimensions as the irreducible representations V^λ

of $\mathfrak{S}_n$, and the same branching rules (see [Mat99], and [GP00] for more details on the Iwahori-Hecke algebras in the general setting of Coxeter groups).

As in the case of the symmetric group $\mathfrak{S}_n$, the character theory of $\mathscr{H}(\mathfrak{S}_n)$ can be encoded in the algebra **Sym** of symmetric functions, *cf.* [Ram91]. A normalized trace on $\mathscr{H}(\mathfrak{S}_n)$ is a function $\tau : \mathscr{H}(\mathfrak{S}_n) \to \mathbb{C}(q)$ that satisfies the two hypotheses:

1. τ is tracial: $\forall a, b, \ \tau(ab) = \tau(ba)$.

2. τ is normalized so that $\tau(1) = 1$.

The condition of positivity is a little more subtle than in the case of the symmetric group. First, for any value of $q \in \mathbb{C}$, we denote $\mathscr{H}_q(\mathfrak{S}_n) = \mathscr{H}(\mathfrak{S}_n) \otimes_{\mathbb{C}(q)} \mathbb{C}$ the specialized Hecke algebra obtained by sending the variable q to its value. This specialized Hecke algebra is a semisimple algebra isomorphic to $\mathbb{C}\mathfrak{S}_n$ (whence with the same representation theory) as soon as $q \neq 0$ and q is not a non-trivial root of unity; in particular, it is the case when $q \in \mathbb{R}_+^*$. Then, assuming $q \in \mathbb{R}_{+*}$, a positive normalized trace, or character on $\mathscr{H}_q(\mathfrak{S}_n)$ is a normalized trace $\mathscr{H}_q(\mathfrak{S}_n) \to \mathbb{C}$ such that (see [KV89]):

3. For all $\sigma_1, \ldots, \sigma_r$ in $\mathfrak{S}_n$, $(\tau(T_{\sigma_i} T_{\sigma_j^{-1}}))_{1 \le i, j \le r}$ is hermitian and non-negative definite.

As in the group case, any trace on $\mathscr{H}(\mathfrak{S}_n)$ is a $\mathbb{C}(q)$-linear combination of the irreducible characters $\zeta^\lambda(q)$ of the modules $V^\lambda(q)$ of $\mathscr{H}(\mathfrak{S}_n)$; we shall also denote by $\chi^\lambda(q)$ the normalized irreducible character of $V^\lambda(q)$. A trace on $\mathscr{H}_q(\mathfrak{S}_n)$ with $q \in \mathbb{R}_+^*$ is positive if and only if it is a positive linear combination of the $\zeta^\lambda(q)$. If $\mu = (\mu_1, \ldots, \mu_r)$ is a partition in $\mathfrak{Y}_n$, let us denote

$$T_\mu = (T_1 \cdots T_{\mu_1 - 1})(T_{\mu_1 + 1} \cdots T_{\mu_1 + \mu_2 - 1}) \cdots (T_{\mu_1 + \cdots + \mu_{r-1} + 1} \cdots T_{\mu_1 + \cdots + \mu_r - 1}).$$

Then, any element $h \in \mathscr{H}(\mathfrak{S}_n)$ is conjugated modulo $[\mathscr{H}(\mathfrak{S}_n), \mathscr{H}(\mathfrak{S}_n)]$ to a unique linear combination of elements T_μ, $\mu \in \mathfrak{Y}_n$ (*cf.* [GP00, §8.2] or [Ram91, §5]). So, a trace $\tau : \mathscr{H}(\mathfrak{S}_n) \to \mathbb{C}(q)$ is entirely determined by the values $\tau(T_\mu)$, and the character theory of $\mathscr{H}(\mathfrak{S}_n)$ is given by the character table

$$\left(\zeta^\lambda(q, \mu) = \zeta^\lambda(q, T_\mu) \right)_{\lambda, \mu \in \mathfrak{Y}_n}.$$

Now, using the notations introduced in §1.2, for any partition μ, we set

$$q_\mu(q, X) = \frac{h_\mu(qX - X)}{(q-1)^{\ell(\mu)}},$$

where $h_\mu = \prod_{i=1}^{\ell(\mu)} s_{\mu_i}$ is the homogeneous symmetric function of type μ; $q_\mu(q, X)$ is an element of **Sym** $\otimes_{\mathbb{C}} \mathbb{C}(q)$. A generalization of the Frobenius formula 1 reads as follows (see [Ram91, Theorem 4.14]):

Proposition 27. *For any partition* $\mu \in \mathfrak{Y}_n$,

$$q_\mu(q, X) = \sum_{\lambda \in \mathfrak{Y}_n} \zeta^\lambda(q, \mu)\, s_\lambda(X).$$

Then, using exactly the same arguments as in the case of $\mathfrak{S}_\infty$, we get the q-analog of Thoma's theorem 3:

Theorem 28. *Suppose* $q \in \mathbb{R}_+^*$. *The set of extremal characters* $\mathcal{X}^*(\mathscr{H}_q(\mathfrak{S}_\infty))$ *can be identified (homeomorphically) with the Thoma simplex* Ω, *and the character* $\chi^\omega(q)$ *corresponding to a point* $\omega = (\alpha, \beta) \in \Omega$ *writes as*

$$\chi^\omega(q, T_\mu) = q_\mu(q, A - B + \gamma E).$$

The decomposition of $\chi^\omega(q)_{|\mathscr{H}_q(\mathfrak{S}_n)}$ *in irreducible characters is*

$$\chi^\omega(q) = \sum_{\lambda \in \mathfrak{Y}_n} \mathbb{P}_{n,\omega}[\lambda]\, \chi^\lambda(q),$$

where $\mathbb{P}_{n,\omega}[\lambda] = (\dim \lambda)\, s_\lambda(A - B + \gamma E)$ *exactly as before.*

This result provides "new" coherent systems of probability measures on $\mathfrak{Y} = \bigsqcup_{n \in \mathbb{N}} \mathfrak{Y}_n$ coming from characters of $\mathscr{H}_q(\mathfrak{S}_\infty)$ instead of characters of $\mathfrak{S}_\infty$. Moreover, letting the parameter q vary in $\mathbb{R}_+^*$, we will obtain one-parameter families of coherent systems. In the following, we will investigate the (asymptotic) behaviour of certain families coming from the so-called Jones-Ocneanu traces of $\mathscr{H}(\mathfrak{S}_\infty)$. The theory developed in the previous paragraphs holds again; however, in the setting of the Hecke algebras, it seems more appriopriate to enounce the algebraic central limit theorem 17 for q-characters instead of characters. The end of this paragraph is devoted to this problem.

For λ and μ partitions, we introduce as in §3.1 the symbol

$$\Sigma_\mu(q, \lambda) = \begin{cases} n(n-1)\ldots(n-|\mu|+1)\, \chi^\lambda(q, \mu \sqcup 1^{n-|\mu|}) & \text{if } n = |\lambda| \geq |\mu|, \\ 0 & \text{otherwise.} \end{cases}$$

At first sight, it is not clear that the functions $\Sigma_\mu(q)$ are observables of diagrams in the algebra $\mathscr{O}$. However, one can expand the symmetric functions $q_\mu(q, X)$ as combinations of power sums $p_\mu(X)$, and therefore write the q-characters as linear combinations of characters, see [Ram91, Theorem 5.4] and [FM10, Proposition 10]:

$$\forall \lambda, \mu, \quad (q-1)^{\ell(\mu)}\, \zeta^\lambda(q, T_\mu) = \sum_{|\nu| = |\mu|} (q^\nu - 1)\, \frac{\langle h_\mu \mid p_\nu \rangle}{\langle p_\nu \mid p_\nu \rangle}\, \zeta^\lambda(\nu)$$

Here, for a partition ν, $q^\nu - 1$ is an abbreviation for $\prod_{i=1}^{\ell(\nu)} q^{\nu_i} - 1$. From this, one deduces that the $\Sigma_\mu(q)$ are indeed observables of diagrams, and that they are related to the Σ_μ by the following relations (notice that the change of basis is triangular):

$$(q-1)^{\ell(\mu)} \, \Sigma_\mu(q) = \sum_{|\nu|=|\mu|} (q^\nu - 1) \, \frac{\langle h_\mu \mid p_\nu \rangle}{\langle p_\nu \mid p_\nu \rangle} \, \Sigma_\nu;$$

$$(q^\mu - 1) \, \Sigma_\mu = \sum_{|\nu|=|\mu|} (q-1)^{\ell(\nu)} \, \langle p_\mu \mid m_\nu \rangle \, \Sigma_\nu(q).$$

Let us fix a parameter $\omega \in \Omega$. Theorem 17 ensures that the variables

$$\Delta_{n,\omega}(\mu) = \sqrt{n} \left(\frac{\Sigma_\mu(\lambda)}{n^{|\mu|}} - p_\mu(\omega) \right)$$

with $\lambda \sim \mathbb{P}_{n,\omega}$ converge jointly to a gaussian vector; since the $\Sigma_\mu(q)$'s are linear combinations of the Σ_μ's, the same result holds for the variables

$$\Delta_{n,\omega}(q,\mu) = \sqrt{n} \left(\frac{\Sigma_\mu(q,\lambda)}{n^{|\mu|}} - q_\mu(q,\omega) \right).$$

We denote by $\Delta_{\infty,\omega}(q,\mu)$ the limiting gaussian variables; their covariances $k(\Delta_{\infty,\omega}(q,\mu^{(1)}), \Delta_{\infty,\omega}(q,\mu^{(2)}))$ are equal by bilinearity to

$$\sum_{|\nu^{(i)}|=|\mu^{(i)}|} C_{\mu^{(1)},\mu^{(2)},\nu^{(1)},\nu^{(2)}}(q) \, k(\Delta_{\infty,\omega}(\nu^{(1)}), \Delta_{\infty,\omega}(\nu^{(2)}))$$

$$= \sum_{\substack{|\nu^{(i)}|=|\mu^{(i)}| \\ a \in \nu^{(1)}, b \in \nu^{(2)}}} ab \, C_{\mu^{(1)},\mu^{(2)},\nu^{(1)},\nu^{(2)}}(q) \, \left(\frac{p_{a+b-1}(\omega)}{p_a(\omega) \, p_b(\omega)} - 1 \right),$$

where

$$C_{\mu^{(1)},\mu^{(2)},\nu^{(1)},\nu^{(2)}}(q) = \frac{(q^{\nu^{(1)} \sqcup \nu^{(2)}} - 1)}{(q-1)^{\ell(\mu^{(1)})+\ell(\mu^{(2)})}} \, \frac{\langle h_{\mu^{(1)}} \mid p_{\nu^{(1)}} \rangle \, \langle h_{\mu^{(2)}} \mid p_{\nu^{(2)}} \rangle}{z_{\nu^{(1)}} \, z_{\nu^{(2)}}}.$$

Theorem 29. *As n goes to infinity, if $\lambda \sim \mathbb{P}_{n,\omega}$, then the variables*

$$\sqrt{n} \left(\chi^\lambda(q,\mu) - \chi^\omega(q,\omega) \right)$$

converge jointly to a gaussian vector, whose covariance matrix is given by the formula above.

In particular, when $\mu^{(1)} = l$ and $\mu^{(2)} = m$ are partitions of length 1, the scalar products $\langle h_{\mu^{(i)}} \mid p_{\nu^{(i)}} \rangle$ are all equal to 1 by Frobenius formula 1, and thus, $k(\Delta_{\infty,\omega}(q,l), \Delta_{\infty,\omega}(q,m))$ is equal to

$$\frac{1}{(q-1)^2} \sum_{\substack{\nu^{(1)} \in \mathfrak{Y}_l \\ \nu^{(2)} \in \mathfrak{Y}_m}} \frac{p_{\nu^{(1)} \sqcup \nu^{(2)}}(q\omega - \omega)}{z_{\nu^{(1)}} \, z_{\nu^{(2)}}} \left(\sum_{\substack{a \in \nu^{(1)} \\ b \in \nu^{(2)}}} ab \left(\frac{p_{a+b-1}(\omega)}{p_a(\omega) \, p_b(\omega)} - 1 \right) \right).$$

In the general case, it does not seem possible to simplify this formula; however, we obtained a simplification in the case of the q-Plancherel measures, see [FM10] and [Mél10].

7.2 Jones-Ocneanu traces and their weights

A **Markov trace** (also-called **Jones-Ocneanu trace**, [Jon87]) on $\mathscr{H}(\mathfrak{S}_n)$ is a normalized trace that satisfies the additional condition:

$$\forall k \in [\![1, n-1]\!], \ \forall a \in \mathscr{H}(\mathfrak{S}_k), \tau(T_k\, a) = \tau(T_k)\,\tau(a) = z\,\tau(a).$$

If z is a fixed parameter, there is a unique corresponding Markov trace on $\mathscr{H}(\mathfrak{S}_n)$, characterized by the values $\tau_z(T_\mu) = z^{|\mu|-\ell(\mu)}$.

Proposition 30. *Let* $X_{q,z} = \frac{[1-q+z]-[z]}{1-[q]}$ *denote the formal alphabet that corresponds to the following specialization of* **Sym***:*

$$\forall k \geq 1, \ p_k(X_{q,z}) = \frac{(1-q+z)^k - z^k}{1-q^k}.$$

The decomposition of τ_z in irreducible q-characters is

$$\tau_z = \sum_{\lambda \in \mathfrak{Y}_n} s_\lambda(X_{q,z})\,\zeta^\lambda(q).$$

Proof. The result is due to Ocneanu (unpublished notes) and Wenzl (see [Wen88]); we give here a very short proof. If $Y_{q,z} = qX_{q,z} - X_{q,z}$, then $p_k(Y_{q,z}) = (q^k - 1)\,p_k(X_{q,z}) = z^k - (1-q+z)^k$. Consequently,

$$P(u) = \sum_{k=1}^\infty \frac{p_k(Y_{q,z})}{k}\,u^k = \log\left(\frac{1-u(1-q+z)}{1-uz}\right)$$

$$H(u) = \sum_{k=0}^\infty h_k(Y_{q,z})\,u^k = \exp P(u) = \frac{1-u(1-q+z)}{1-uz}.$$

Hence, $h_k(Y_{q,z}) = (q-1)z^{k-1}$, and

$$q_\mu(q, X_{q,z}) = \frac{h_\mu(Y_{q,z})}{(q-1)^{\ell(\mu)}} = \prod_{i=1}^{\ell(\mu)} z^{\mu_i-1} = z^{|\mu|-\ell(\mu)} = \tau_z(T_\mu).$$

Now, because of the q-Frobenius formula 27, for any partition μ,

$$\tau_z(T_\mu) = q_\mu(q, X_{q,z}) = \sum_{\lambda \in \mathfrak{Y}_n} s_\lambda(X_{q,z})\,\zeta^\lambda(q, T_\mu).$$

Since a trace is characterized by its values on the T_μ's, the result follows immediately. $\square$

There is a hook length formula for the specializations $s_\lambda(X_{q,z})$ of the Schur functions, see for instance [Mac95, §1.3]:

$$s_\lambda(X_{q,z}) = q^{n(\lambda)} \prod_{\square \in \lambda} \frac{(1-q) + z(1-q^{c(\square)})}{1 - q^{h(\square)}}$$

where $n(\lambda) = \sum_{i=1}^{\ell(\lambda)}(i-1)\lambda_i$, and if $\square = (i,j)$ is a box of the Young diagram of λ, $c(\square) = i - j$ is the **content** of $\square$. Now, let us make the change of variables:

$$z = -(1-q)(1-t).$$

Then, assuming that $q \in \mathbb{R}_+^*$ and $t \in [0,1]$, one sees that $s_\lambda(X_{q,z})$ is positive, because

$$s_\lambda(X_{q,z}) = \frac{q^{b(\lambda)}}{\prod_{\square \in \lambda}\{h(\square)\}_q} \left(\prod_{\square \in \lambda} t + (1-t)q^{c(\square)} \right)$$

with $\{i\}_q = \frac{1-q^i}{1-q}$ for any positive integer i. Another way to see that $s_\lambda(X_{q,z})$ is positive in this situation is to notice that

$$X_{q,z} = A - \overline{B},$$

where A and B are the alphabets associated to the point of the Thoma simplex $\omega_{q,t} = (\alpha_{q,t}, \beta_{q,t})$, with

$$\alpha_{q,t} = \frac{[t(1-q)]}{1-[q]} = (t(1-q), t(1-q)q, t(1-q)q^2, \ldots);$$

$$\beta_{q,t} = \frac{[(1-t)(1-q)]}{1-[q]} = ((1-t)(1-q), (1-t)(1-q)q, (1-t)(1-q)q^2, \ldots).$$

In particular:

Proposition 31. *If $z = -(1-q)(1-t)$ with $t \in [0,1]$ and $q \in \mathbb{R}_+^*$, then the Markov trace τ_z is an irreducible (positive) character of $\mathscr{H}_q(\mathfrak{S}_\infty)$, and it is equal to $\chi^{\omega_{q,t}}$ with $\omega_{q,t}$ defined by the formulas above.*

Example. Suppose that $t = 1$. Then, $z = 0$ and $\tau = \tau_0$ is the **regular trace** of $\mathscr{H}_q(\mathfrak{S}_\infty)$:

$$\forall \omega \in \mathfrak{S}_\infty, \ \ \tau(T_\omega) = \begin{cases} 1 & \text{if } \omega = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Although the definition of the regular does not seem to depend on the value of q, the weights of τ in its decomposition in irreducible characters vary with q:

$$\forall n, \ \ \tau_{|\mathscr{H}_q(\mathfrak{S}_n)} = (1-q)^n \sum_{\lambda \in \mathfrak{Y}_n} (\dim \lambda) \, s_\lambda(1, q, q^2, \ldots) \, \chi^\lambda(q).$$

The probability measures corresponding to this decomposition have been studied extensively in [FM10]. In the following, we generalize the results of this paper by using the general theorems proved in Sections 4 and 5.

7.3 (q, t)-Plancherel measures and their asymptotics

For $q \in \mathbb{R}_+^*$ and $t \in [0, 1]$, we call **(q,t)-Plancherel measure** of size n the probability measure $\mathbb{P}_{n,q,t}$ on $\mathfrak{Y}_n$ corresponding to the decomposition of the Markov trace τ_z with $z = -(1-q)(1-t)$. In other words, $\mathbb{P}_{n,q,t} = \mathbb{P}_{n,\omega_{q,t}}$, and

$$\mathbb{P}_{n,q,t}[\lambda] = (1-q)^n (\dim \lambda) \, s_\lambda \left(\frac{[t] - [t-1]}{1 - [q]} \right).$$

The (q, t)-Plancherel measures have the following symmetry:

$$\mathbb{P}_{n,q,t}[\lambda] = \mathbb{P}_{n,q^{-1},t}[\lambda'].$$

Hence, we shall always suppose that $q \in \,]0, 1[$, the other case following by a symmetry.

Remark. When q goes to 1, one recovers the usual Plancherel measures defined on page 11

$$\mathbb{P}_{n,q=1,t}[\lambda] = \mathbb{P}_n[\lambda] = \frac{(\dim \lambda)^2}{n!}.$$

Hence, the (q, t)-Plancherel measures form a two-parameter deformation of the coherent system of Plancherel measures. They are also a generalization of the q-Plancherel measures defined in [Ker92, Str08] and studied in [FM10].

Using the results of §2, one can interpret the (q, t)-Plancherel measures as push-forwards of probability measures on the symmetric groups. Let us detail a little bit these measures $\mathbb{Q}_{n,q,t}$ such that $\Lambda_\star \mathbb{Q}_{n,q,t} = \mathbb{P}_{n,q,t}$. We could of course take $\mathbb{Q}_{n,q,t}[\sigma] = \mathbb{Q}_{n,\omega_{q,t}}[\sigma]$, but it will be a little bit easier[13] to describe the probability measures

$$\mathbb{Q}_{n,q,t}[\sigma] = \mathbb{Q}_{n,\omega_{q,t}}[\sigma^{-1}] = F_\sigma(\omega_{q,t}) = L_{c(\sigma)}(\omega_{q,t}).$$

Taking σ^{-1} instead of σ does not change the fact that $\Lambda_\star \mathbb{Q}_{n,q,t} = \mathbb{P}_{n,q,t}$, because the bilateral Kazhdan-Lusztig cells $\{\sigma \in \mathfrak{S}_n \mid \Lambda(\sigma) = \lambda\}$ are stabilized by the inversion $\sigma \mapsto \sigma^{-1}$. Now, for any composition $c \in \mathfrak{C}_n$, if $\mathrm{comaj}(c) = \sum_{d \in D(c)} n - d$, then

$$\forall q \in \,]0, 1[, \quad L_c(1, q, q^2, \ldots) = \frac{q^{\mathrm{comaj}(c)}}{(1-q)(1-q^2)\cdots(1-q^n)},$$

---

[13] In return, we lose the direct interpretation of the random permutations as generalized shuffles.

see [Sta91, §7.19]. Using the definition of the specialization $L_c(\omega)$ (*cf.* Paragraph 2.2), one sees that:

$$\mathbb{Q}_{n,q,t}[\sigma] = L_{c(\sigma)}(\omega_{q,t}) = \sum_{i=0}^{n} L_{c(\sigma)_{[\![1,i]\!]}}(\alpha_{q,t})\, L_{\overline{c(\sigma)_{[\![i+1,n]\!]}}}(\beta_{q,t})$$

$$= \sum_{i=0}^{n} L_{c(\sigma)_{[\![1,i]\!]}}(\alpha_{q,t})\, L_{c(\sigma\omega_0)_{[\![1,n-i]\!]}}(\beta_{q,t})$$

$$= \sum_{i=0}^{n} \frac{t^i\,(1-t)^{n-i}}{\{i!\}_q\,\{n-i!\}_q}\, q^{\mathrm{comaj}(c(\sigma)_{[\![1,i]\!]})+\mathrm{comaj}(c(\sigma\omega_0)_{[\![1,n-i]\!]})}.$$

For w word of size n, we set $\mathbb{Q}_{n,q}[w] = \frac{q^{\mathrm{comaj}(w)}}{\{n!\}_q}$; for any $q \in \mathbb{R}_+^*$, $\mathbb{Q}_{n,q}$ is a probability measure on $\mathfrak{S}_n$.

Proposition 32. *As t varies in $[0,1]$, the measures $\mathbb{Q}_{n,q,t}$ provide an interpolation between $\mathbb{Q}_{n,q}[\sigma]$ and $\mathbb{Q}_{n,q}[\sigma\omega_0]$:*

$$\mathbb{Q}_{n,q,t}[\sigma] = \sum_{i=0}^{n} t^i\,(1-t)^{n-i}\,\mathbb{Q}_{i,q}[\sigma_{[\![1,i]\!]}]\,\mathbb{Q}_{n-i,q}[(\sigma\omega_0)_{[\![1,n-i]\!]}]$$

where $\sigma_{[\![1,i]\!]}$ denotes the word $\sigma(1)\sigma(2)\cdots\sigma(i)$, and similarly for $(\sigma\omega_0)_{[\![1,n-i]\!]}$.

It was *a priori* a non trivial fact that these interpolations correspond *via* RSK to the weights of the Markov traces $\tau_{-(1-q)(1-t)}$. On the other hand, let us notice that the parts $\mathbb{Q}_{i,q}[\sigma_{[\![1,i]\!]}]\,\mathbb{Q}_{n-i,q}[(\sigma\omega_0)_{[\![1,n-i]\!]}]$ of $\mathbb{Q}_{n,q,t}[\sigma]$ have a simple geometric interpretation. If σ is a permutation, let us draw its corresponding ribbon, see §2.1. We draw a line after the i-th box, and we define $m_i(\sigma)$ as the sum of the following quantities:

- for each box $\boxed{\vee}$ before the line and with another box just under, the distance between $\boxed{\vee}$ and the line in the ribbon;
- for each box $\boxed{<}$ after the line and with another box just at the left, the distance between $\boxed{<}$ and the line in the ribbon.

Then, the i-th weight is simply $\frac{q^{m_i(\sigma)}}{\{i!\}_q\,\{n-i!\}_q}$; for instance, if $i = 5$ and $\sigma = 64182357$, then $m_i(\sigma) = (4_6 + 3_4 + 1_8) + (0_3 + 1_5 + 2_7) = 11$.

Now, let us describe the asymptotics of the (q,t)-Plancherel measures. First, we give the algebraic central limit theorem, noticing that

$$p_k(\omega_{q,t}) = \frac{(t^k - (t-1)^k)(1-q)^k}{1-q^k}$$

for any $k \geq 1$.

Proposition 33. *If $\lambda \sim \mathbb{P}_{n,q,t}$, then as n goes to infinity, the random variables*

$$\sqrt{n}\left(\chi^\lambda(q, T_l) - (t-1)^{l-1}(1-q)^{l-1}\right)$$

converge jointly to a gaussian vector $(\Delta_{\infty,q,t}(q,l))_{l\geq 1}$ with covariances

$$(1-q)^{-l-m+2}\, k(\Delta_{\infty,q,t}(q,l), \Delta_{\infty,q,t}(q,m))$$
$$= \sum_{\substack{\nu^{(1)}\in\mathfrak{Y}_l \\ \nu^{(2)}\in\mathfrak{Y}_m \\ a\in\nu^{(1)},\, b\in\nu^{(2)}}} \frac{(t-1)^\nu - t^\nu}{z_{\nu^{(1)}}\, z_{\nu^{(2)}}} \left(\frac{1-u^{a+b-1}}{1-q^{a+b-1}}\frac{1-u}{1-q}\frac{1-q^a}{1-u^a}\frac{1-q^b}{1-u^b} - 1\right)$$

where $u = \frac{t-1}{t}$, $\nu = \nu^{(1)} \sqcup \nu^{(2)}$, and $(t-1)^\nu - t^\nu = \prod_{i=1}^{\ell(\nu)}((t-1)^{\nu_i} - t^{\nu_i})$.

Remark. When $t = 1$, this formula can be simplified by using a Möbius inversion formula, and it becomes

$$k(\Delta_{\infty,q,t}(q,l), \Delta_{\infty,q,t}(q,m)) = \frac{(q-q^2)^{l+m-3}(1-q^2)\{l-1\}_q\{m-1\}_q}{\{l+m-1\}_q\{l+m-2\}_q\{l+m-3\}_q}$$

see [Mél10].

Now, let us give the geometric version of the central limit theorem. We have drawn below a random Young diagram taken according to the (q,t)-Plancherel measure of parameters $n = 300$, $q = \frac{1}{2}$ and $t = \frac{2}{3}$. With $q = \frac{1}{2}$,
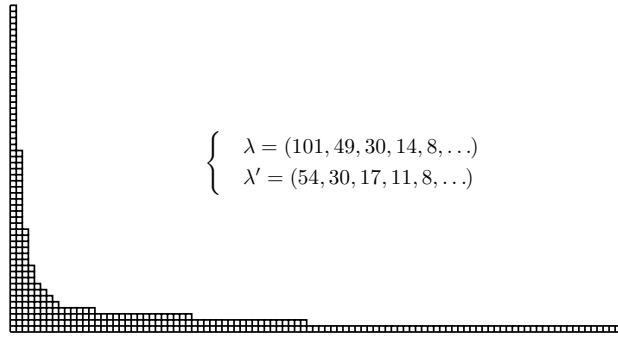

$$\begin{cases} \lambda = (101, 49, 30, 14, 8, \ldots) \\ \lambda' = (54, 30, 17, 11, 8, \ldots) \end{cases}$$

Fig. 8 A random Young diagram under a (q,t)-Plancherel measure.

it is not hard to see that the first rows and the first columns seem to follow geometric progressions of reason $\frac{1}{2}$. Moreover, with $t = \frac{2}{3}$, the ratio between the first row and the first column seems to be close to $\frac{t}{1-t} = 2$. This is in adequation with Theorems 15 and 19 that state the following (notice that we are in the case of inequal coordinates):

Proposition 34. *If $\lambda \sim \mathbb{P}_{n,q,t}$, then*

$$\lambda_i = n\, t(1-q)q^{i-1} + n^{1/2}\, X_i + o(n^{1/2});$$
$$\lambda'_j = n\, (1-t)(1-q)q^{j-1} + n^{1/2}\, Y_j + o(n^{1/2}),$$

where $(X_i, Y_j)_{i \geq 1, j \geq 1}$ is a gaussian vector with covariances

$$k(X_i, X_j) = \delta_{i,j}\, t(1-q)q^{i-1} - t^2(1-q)^2 q^{i+j-2};$$
$$k(Y_i, Y_j) = \delta_{i,j}\, (1-t)(1-q)q^{i-1} - (1-t)^2(1-q)^2 q^{i+j-2};$$
$$k(X_i, Y_j) = -t(1-t)(1-q)^2 q^{i+j-2}.$$

In particular, for $t = 1$, we recover Theorems 1 and 2 of [FM10]. Proposition 34 also gives the asymptotic behaviour of the length of a longest increasing or decreasing subword in a random permutation taken according to the probability measure described in Proposition 32.

Conclusion

Let us summarize the results obtained in this paper:

1. To any irreducible character of $\mathfrak{S}_\infty$ or of $\mathscr{H}_q(\mathfrak{S}_\infty)$ with $q \in \mathbb{R}_+^*$, we have associated a point ω in the Thoma simplex Ω, and two random models: a model $(\mathbb{Q}_{n,\omega})_{n \in \mathbb{N}}$ of random permutations obtained by generalized riffle shuffles, and a model $(\mathbb{P}_{n,\omega})_{n \in \mathbb{N}}$ of random partitions related to the restriction of the irreducible character to the finite symmetric groups or their Hecke algebras.

2. The computation of the probabilities of both models can be done using specializations of combinatorial Hopf algebras: the Hopf algebra **FQSym** for the permutations, and the Hopf algebra **Sym** for the partitions. This interpretation enabled us to see that the two models are related by the RSK correspondence.

3. The model of random partitions satisfies a law of large numbers and a central limit theorem. In both cases, the asymptotic result can be given an algebraic flavour (for the values of the characters), or a geometric flavour (for the first rows and columns of the partitions). The geometric central limit theorem is much more involved than the other results, and cannot be proved by using only techniques of moments. Hence, one had to relate random permutations and partitions to some conditioned random walks, and to the behaviour of the eigenvalues of brownian hermitian matrices.

When $\omega = \mathbf{0}$, our methods do not describe with sufficient precision the asymptotic behaviour of the random partitions; however, they can be adapted to this purpose, see [IO02]. One obtains again a law of large numbers and a central limit theorem, and this CLT seems to have a universality property; *cf.* [Mél11]. Hence, the next step in the asymptotic study of the models presented in this article would be to understand what happens "in the neighbourhood" of the point $\mathbf{0}$ of the Thoma simplex. More precisely, we would like to understand what happens when ω *varies with n and converges to* $\mathbf{0}$. Let us give a concrete example — this is one of the two cases studied in [Mél11], and we refer also to [Bia01]. Suppose that

$$\omega_n = \left(\left(\frac{1}{N}, \ldots, \frac{1}{N}, 0, 0, \ldots \right), (0, 0, \ldots) \right),$$

with N coordinates equal to $\frac{1}{N}$, and $cN \simeq \sqrt{n}$. Then, there is a limit shape for the random partitions $\lambda \sim \mathbb{P}_{n,\omega_n}$; it has been found by P. Biane in [Bia01], and it is drawn on Figure 9 in the case when $c = 1$.
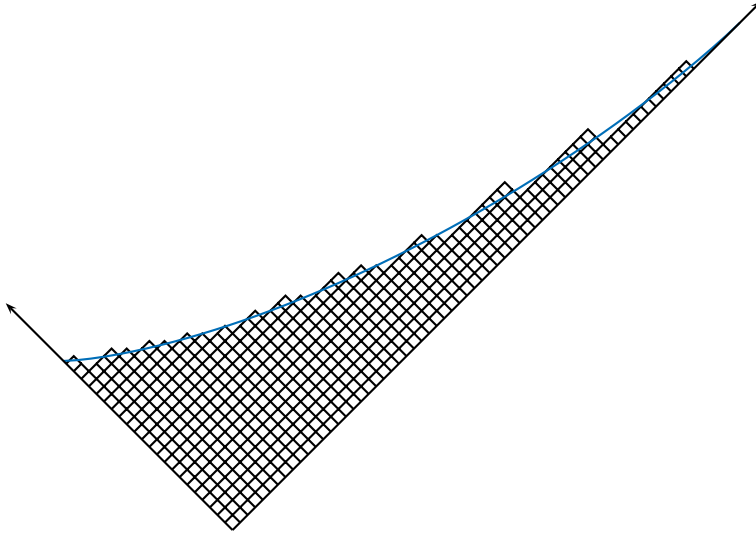

Fig. 9 Random partition under the Schur-Weyl measure of parameter $c = 1$.

Now, the interesting fact is that the fluctuations of the partitions under these measures (called **Schur-Weyl measures**) are described by a gaussian free field that does not depend on c (or, to be more precise, only by a translation of the fluctuations on the x-axis). More generally, we conjecture the following (the statement is willingly left unprecise):

Conjecture 35. *This asymptotic behaviour is universal among models of random partitions associated to parameters $(\omega_n)_{n\in\mathbb{N}}$ converging to $\mathbf{0}$ at speed $1/\sqrt{n}$ as n goes to infinity.*

There is one example whose study would be very interesting in order to prove this conjecture — it has been suggested by M. Bożejko. In some sense, the Schur-Weyl measures provide another one-parameter deformation of the Plancherel measures: indeed, when c goes to 0, the limit shape of the partitions under Schur-Weyl measures converges to the one corresponding to the Plancherel measures. On the other hand, the limit shapes under Schur-Weyl measures correspond *via* the Markov-Krein correspondence (see [Ker98]) to the Marčenko-Pastur measures, that are a one-parameter deformation of the Wigner semicircle law (corresponding *via* Markov-Krein to the limit shape of the random partitions under Plancherel measures). Now, the **free Meixner laws** (see *e.g.* [BB06]) form a two-parameter deformation of the Wigner measure, and they generalize the Marčenko-Pastur laws. So it would be extremely interesting to construct models of random partitions corresponding to these laws, and to study their fluctuations. One can hope that with two parameters, the computations will be sufficiently generic in order to understand the behaviour of random partitions in the general situation of Conjecture 35. The results of this paper have been obtained in a similar way: starting from the q-Plancherel measures already studied in [FM10], we have gone to the case of (q,t)-Plancherel measures, and we have then realized that our results hold in fact for any model of random partitions associated to a point of the Thoma simplex.

References

AGZ09. G. W. Anderson, A. Guionnet, and O. Zeitouni. *An Introduction to Random Matrices*, volume 118 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 2009.

AS05. M. Aguiar and F. Sottile. Structure of the Malvenuto-Reutenauer Hopf algebra of permutations. *Adv. Math.*, 191(2):225–275, 2005.

BB06. M. Bożejko and W. Bryc. On a class of free Lévy laws related to a regression problem. *J. Funct. Anal.*, pages 59–77, 2006.

BBO05. P. Biane, P. Bougerol, and N. O'Connell. Littelmann paths and brownian paths. *Duke Math. J.*, 130(1):127–167, 2005.

Bia01. P. Biane. Approximate factorization and concentration for characters of symmetric groups. *Internat. Math. Res. Notices*, 4:179–192, 2001.

Bia09. P. Biane. From Pitman's theorem to crystals. *Advanced Studies in Pure Mathematics*, 55:1–13, 2009.

Bil69. P. Billingsley. *Convergence of Probability Measures*. Wiley, 1969.

Bri69. D. Brillinger. The calculation of cumulants via conditioning. *Ann. Inst. Statist. Math.*, 21:375–390, 1969.

DFŚ10. M. Dołega, V. Féray, and P. Śniady. Explicit combinatorial interpretation of Kerov character polynomials as numbers of permutation factorizations. *Adv. Math.*, 225(1):81–120, 2010.

DHT01.   G. Duchamp, F. Hivert, and J.-Y. Thibon. Noncommutative symmetric functions VI: Free quasi-symmetric functions and related algebras. `arXiv:math/0105065v1`, 2001.

Dys62.   F. J. Dyson. A brownian-motion model for the eigenvalues of a random matrix. *J. Math. Phys.*, 3:1191–1198, 1962.

FM10.   V. Féray and P.-L. Méliot. Asymptotics of q-Plancherel measures. To appear in *Probability Theory and Related Fields*, `arXiv:1001.2180v1 [math.RT]`, 2010.

Ful97.   W. Fulton. *Young Tableaux with Applications to Representation Theory and Geometry*, volume 35 of *London Mathematical Society Students Texts*. Cambridge University Press, 1997.

Ful02.   J. Fulman. Applications of symmetric functions to cycle and increasing subsequence structure after shuffles. *J. Alg. Combin.*, 16:165–194, 2002.

Ges94.   I. M. Gessel. Multipartite P-partitions and inner products of skew Schur functions. In *Combinatorics and algebra (Boulder, Colo., 1983)*, pages 289–317. Amer. Math. Soc., 1994.

GK10.   R. Gohm and C. Köstler. Noncommutative independence from characters of the infinite symmetric group $s\infty$. `arXiv:1005.5726`, 2010.

GP00.   M. Geck and G. Pfeiffer. *Characters of Finite Coxeter Groups and Iwahori-Hecke Algebras*, volume 21 of *London Mathematical Society Monographs*. Oxford University Press, 2000.

IK99.   V. Ivanov and S. Kerov. The algebra of conjugacy classes in symmetric groups, and partial permutations. In *Representation Theory, Dynamical Systems, Combinatorial and Algorithmical Methods III*, volume 256 of *Zapiski Nauchnyh Seminarov POMI*, pages 95–120, 1999.

IO02.   V. Ivanov and G. Olshanski. Kerov's central limit theorem for the Plancherel measure on Young diagrams. In *Symmetric Functions 2001: Surveys of Developments and Perspectives*, volume 74 of *NATO Science Series II. Mathematics, Physics and Chemistry*, pages 93–151, 2002.

JK81.   G. D. James and A. Kerber. *The Representation Theory of the Symmetric Group*, volume 16 of *Encyclopedia of Mathematics and its Applications*. Addison-Wesley, 1981.

Jon87.   V. F. R. Jones. Hecke algebra representations of braid groups and link polynomials. *Ann. of Math.*, 126(2):335–388, 1987.

Ker92.   S. V. Kerov. q-analogue of the hook walk algorithm and random Young tableaux. *Funct. Anal. Appl.*, 26(3):179–187, 1992.

Ker93.   S. V. Kerov. Gaussian limit for the Plancherel measure of the symmetric group. *Comptes Rendus Acad. Sci. Paris, Série I*, 316:303–308, 1993.

Ker98.   S. V. Kerov. Interlacing measures. *Amer. Math. Soc. Transl.*, 181(2):35–83, 1998.

KO94.   S. V. Kerov and G. Olshanski. Polynomial functions on the set of Young diagrams. *Comptes Rend. Acad. Sci. Paris, Série I*, 319:121–126, 1994.

KOV04.   S. V. Kerov, G. Olshanski, and A. M. Vershik. Harmonic analysis on the infinite symmetric group. *Invent. Math.*, 158:551–642, 2004.

KV77.   S. V. Kerov and A. M. Vershik. Asymptotics of the Plancherel measure of the symmetric group and the limiting form of Young tableaux. *Doklady AN SSSR*, 233(6):1024–1027, 1977.

KV81a.   S. V. Kerov and A. M. Vershik. Asymptotic theory of characters of the symmetric group. *Funct. Anal. Appl.*, 15(4):246–255, 1981.

KV81b.   S. V. Kerov and A. M. Vershik. Characters and factor representations of the infinite symmetric group. *Soviet Math. Doklady*, 23:389–392, 1981.

KV84.   S. V. Kerov and A. M. Vershik. Characters, factor representations and K-functor of the infinite symmetric group. In *Proc. Intern. Conf. on Operator Algebras and Group Representations 1980*, volume 2 of *Monographs and Studies in Mathematics*, pages 23–32. Pitman, London, 1984.

KV89.   S. V. Kerov and A. M. Vershik. Characters and realizations of representations of an infinite-dimensional Hecke algebra, and knot invariants. *Soviet Math. Doklady*, 38(1):134–137, 1989.

Lit95.   P. Littelmann. Paths and root operators in representation theory. *Ann. Math.*, 142(3):499–525, 1995.

LLT02.   A. Lascoux, B. Leclerc, and J.-Y. Thibon. The plactic monoid. In *Algebraic Combinatorics on Words*. Cambridge University Press, 2002.

LS77.   B. F. Logan and L. A. Shepp. A variational problem for random Young tableaux. *Adv. Math.*, 26:206–222, 1977.

Mac95.   I. G. Macdonald. *Symmetric functions and Hall polynomials*. Oxford Mathematical Monographs. Oxford University Press, 2nd edition, 1995.

Mat99.   A. Mathas. *Iwahori-Hecke algebras and Schur algebras of the symmetric group*, volume 15 of *University Lecture Series*. Amer. Math. Soc., 1999.

Mél10.   P.-L. Méliot. Gaussian concentration of the q-characters of the Hecke algebras of type A. `arXiv:1009.4288v1 [math.RT]`, 2010.

Mél11.   P.-L. Méliot. Kerov's central limit theorem for Schur-Weyl and Gelfand measures. To appear in the proceedings of the 23th International Conference on Formal Power Series and Algebraic Combinatorics (Reykjavik), 2011.

MR95.   C. Malvenuto and C. Reutenauer. Duality between quasi-symmetric functions and Solomon descent algebra. *J. Algebra*, 177:967–982, 1995.

O'C03.   N. O'Connell. A path-transformation for random walks and the Robinson-Schensted correspondence. *Trans. Amer. Math. Soc.*, 355:3669–3697, 2003.

Oko97.   A. Okounkov. On representations of the infinite symmetric group. Representation theory, dynamical systems, combinatorial and algorithmic methods II (A. M. Vershik, ed.). *Zapiski Seminarov POMI*, 240:167–229, 1997.

OV04.   A. Okounkov and A. M. Vershik. A new approach to the representation theory of the symmetric groups, II. *Zapiski Seminarov POMI*, 307:57–98, 2004.

OY02.   N. O'Connell and M. Yor. A representation for non-colliding random walks. *Electron. Comm. Probab.*, 7:1–12, 2002.

Pit75.   J. W. Pitman. One-dimensional Brownian motion and the three-dimensional Bessel process. *Advances in Applied Probability*, 7:511–526, 1975.

Ram91.   A. Ram. A Frobenius formula for the characters of the Hecke algebras. *Invent. Math.*, 106:461–488, 1991.

Rob38.   G. Robinson. On the representations of the symmetric group. *Amer. J. Math.*, 60(3):745–760, 1938.

RY91.   D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*, volume 293 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, 3rd edition, 1991.

Sch61.   C. Schensted. Longest increasing and decreasing subsequences. *Canadian Journal of Mathematics*, 13:179–191, 1961.

Śni06a. P. Śniady. Asymptotics of characters of symmetric groups, genus expansion and free probability. *Discrete Math.*, 306(7):624–665, 2006.

Śni06b. P. Śniady. Gaussian fluctuations of characters of symmetric groups and of Young diagrams. *Probability Theory and Related Fields*, 136(2):263–297, 2006.

Sta91.   R. P. Stanley. *Enumerative combinatorics*. Cambridge University Press, 1991.

Sta01.   R. P. Stanley. A generalized riffle shuffle and quasisymmetric functions. *Annals of Combinatorics*, 5:479–491, 2001.

Str08.   E. Strahov. A differential model for the deformation of the Plancherel growth process. *Adv. Math*, 217(6):2625–2663, 2008.

Tha87.   D. H. Thang. Random operators in Banach spaces. *Probability and Mathematical Statistics*, 8:155–167, 1987.

Tho64.   E. Thoma. Die unzerlegbaren, positiv-definiten Klassenfunktionen der abzählbar unendlichen symmetrischen Gruppe. *Math. Zeitschrift*, 85:40–61, 1964.

VV95.   M. V. Velasco and A. R. Villena. A random Banach-Steinhaus theorem. *Proc. Amer. Math. Soc.*, 123(8):2489–2497, 1995.

Was81.   A. J. Wassermann. Automorphic actions of compact groups on operator algebras. Ph.D. Thesis, University of Pennsylvania, 1981.

Wen88.   H. Wenzl. Hecke algebras of type A_n and subfactors. *Invent. Math.*, 92:349–383, 1988.

Zel81.   A. Zelevinsky. *Representations of finite classical groups: A Hopf algebra approach*, volume 869 of *Lecture Notes in Mathematics*. Springer-Verlag, 1981.